Global SMEFT fits guided by ML

JHEP 03 (2023) 033

2211.02058

1st COMETA General Meeting

29/02/24

Izmir

Jaco ter Hoeve

VU Amsterdam & Nikhef Theory Group

Outline

Mapping the SMEFT with SMEFiT

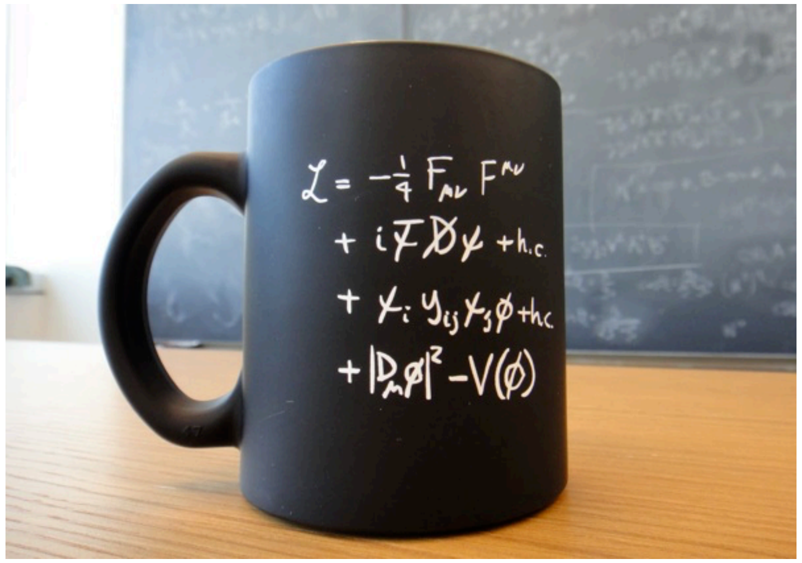
ML assisted observables: the ML4EFT framework

Case study: ML observables in the Higgs and top sector

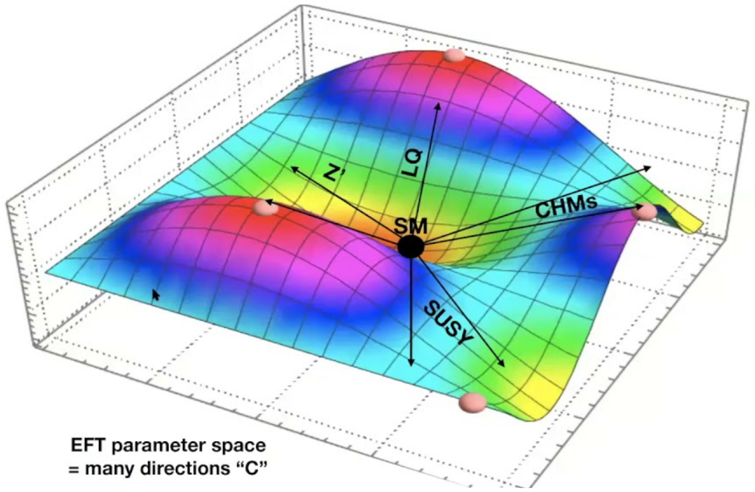
A combined ML4EFT + SMEFiT analysis

The Standard Model as an EFT

$$\mathscr{L}_{\mathrm{SMEFT}} = \mathscr{L}_{\mathrm{SM}} + \sum_i^{N_{d5}} \frac{c_i}{\Lambda}\mathcal{O}_i^{(5)} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_i^{N_{d7}} \frac{c_i}{\Lambda^3}\mathcal{O}_i^{(7)} + \sum_i^{N_{d8}} \frac{b_i}{\Lambda^4}\mathcal{O}_i^{(8)} + \ldots$$



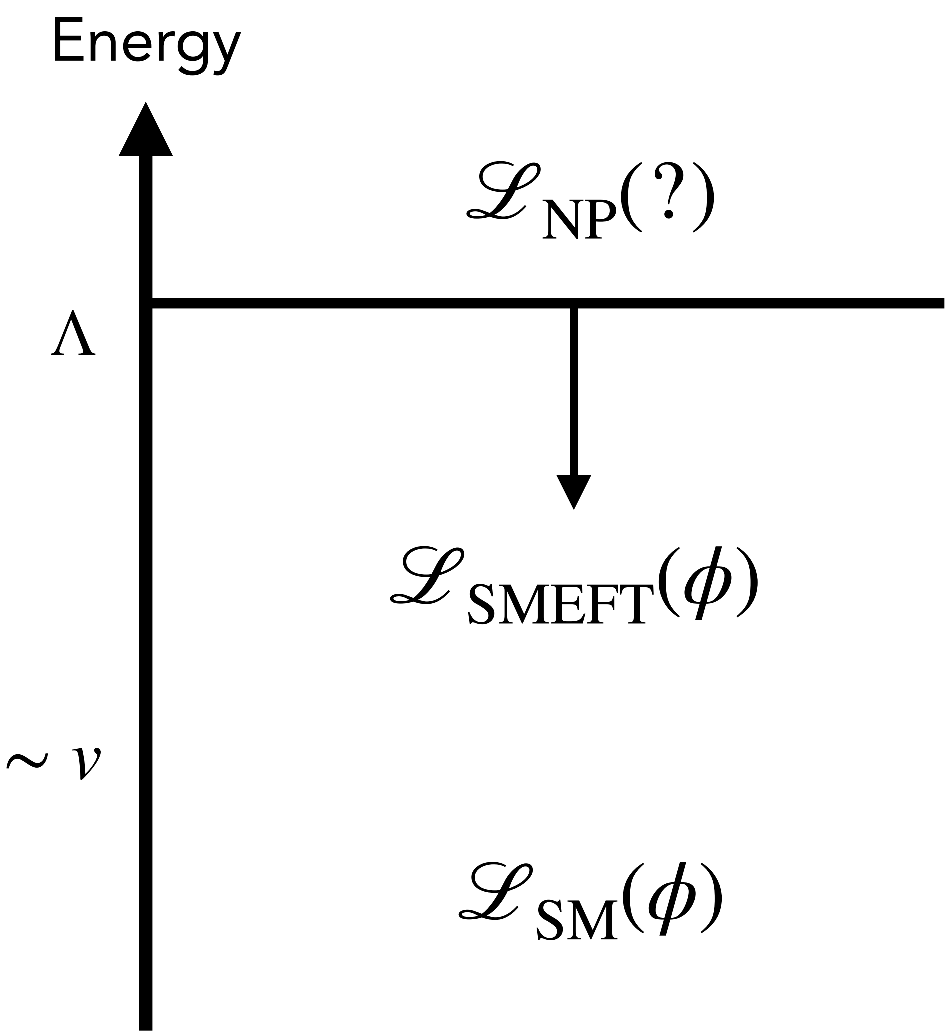
+

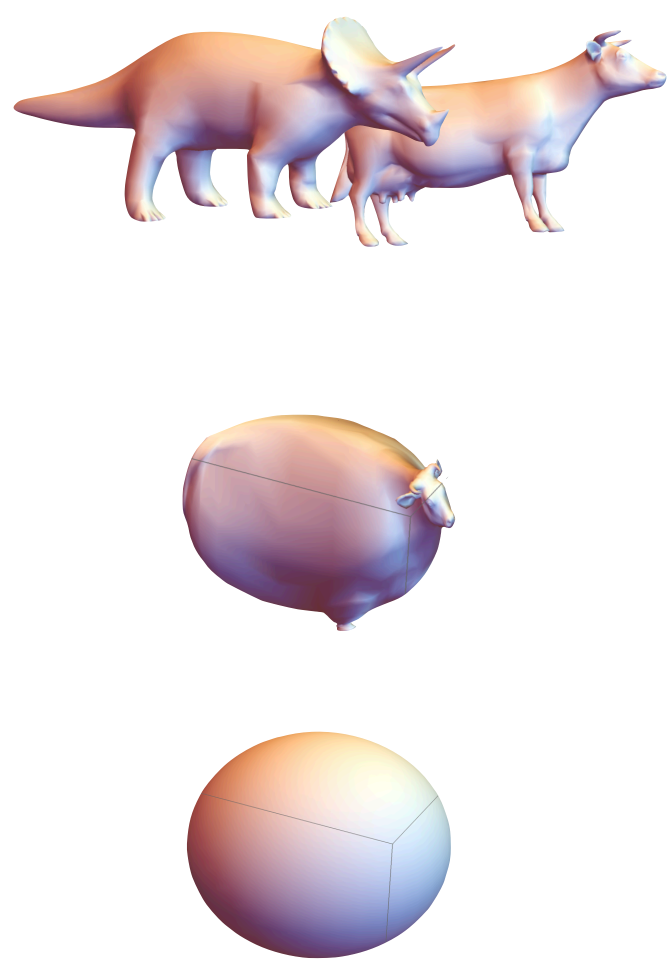


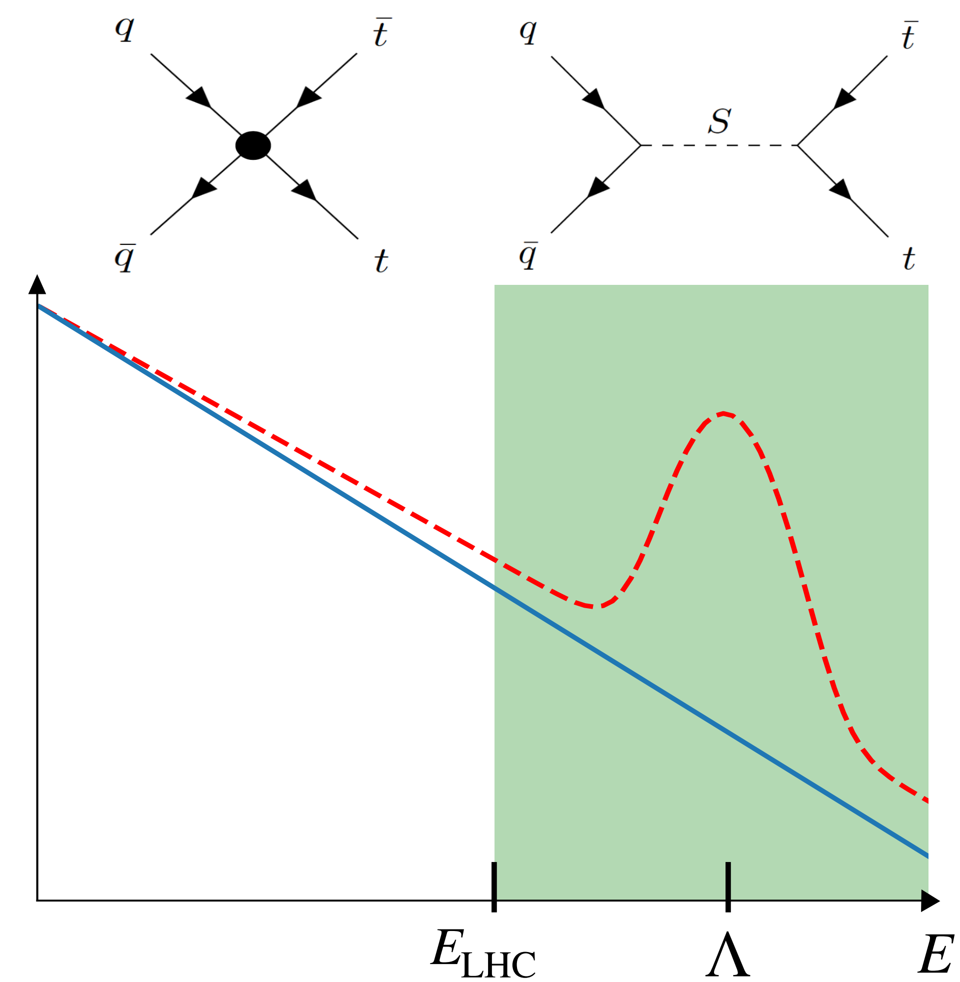
- **Low energy limit** of generic UV-complete theories at high energies
- Assumes the **SM field content and symmetries**
- **Complete basis** at any given mass dimension

The Standard Model as an EFT

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \sum_i^{N_{d5}} \frac{c_i}{\Lambda}\mathcal{O}_i^{(5)} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_i^{N_{d7}} \frac{c_i}{\Lambda^3}\mathcal{O}_i^{(7)} + \sum_i^{N_{d8}} \frac{b_i}{\Lambda^4}\mathcal{O}_i^{(8)} + \ldots$$

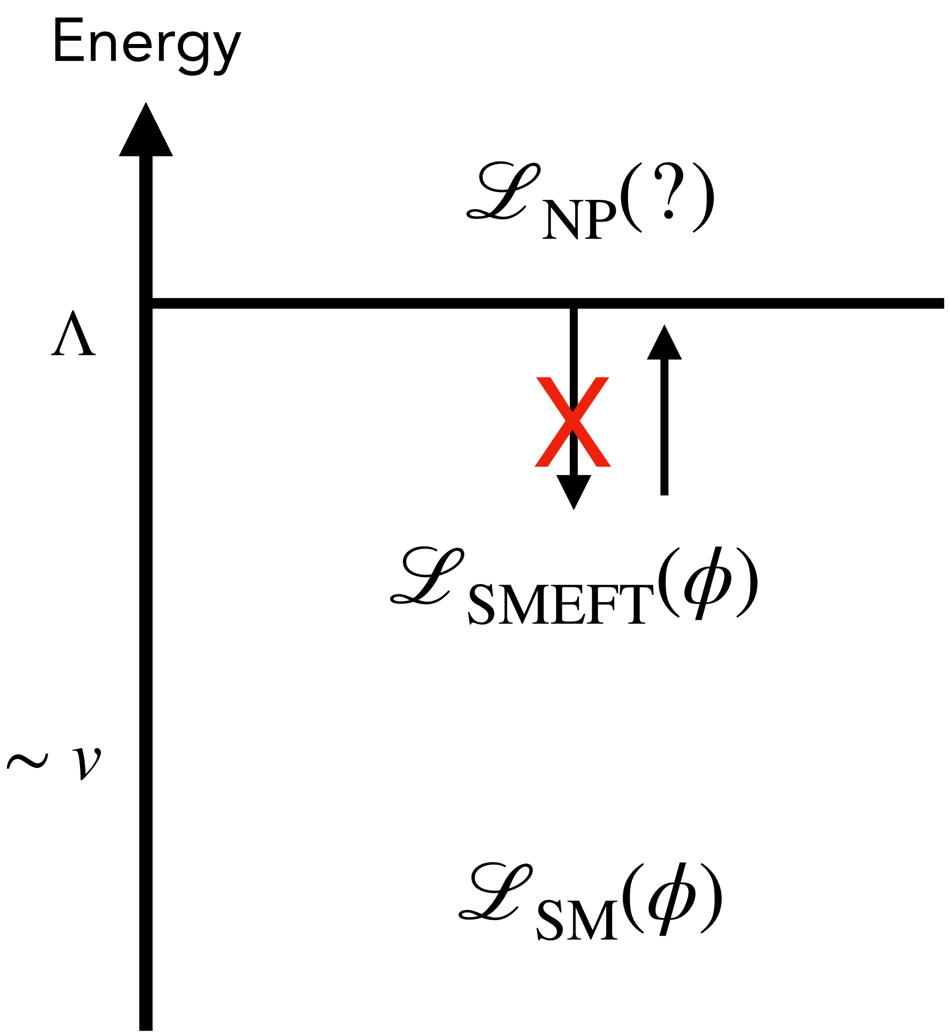


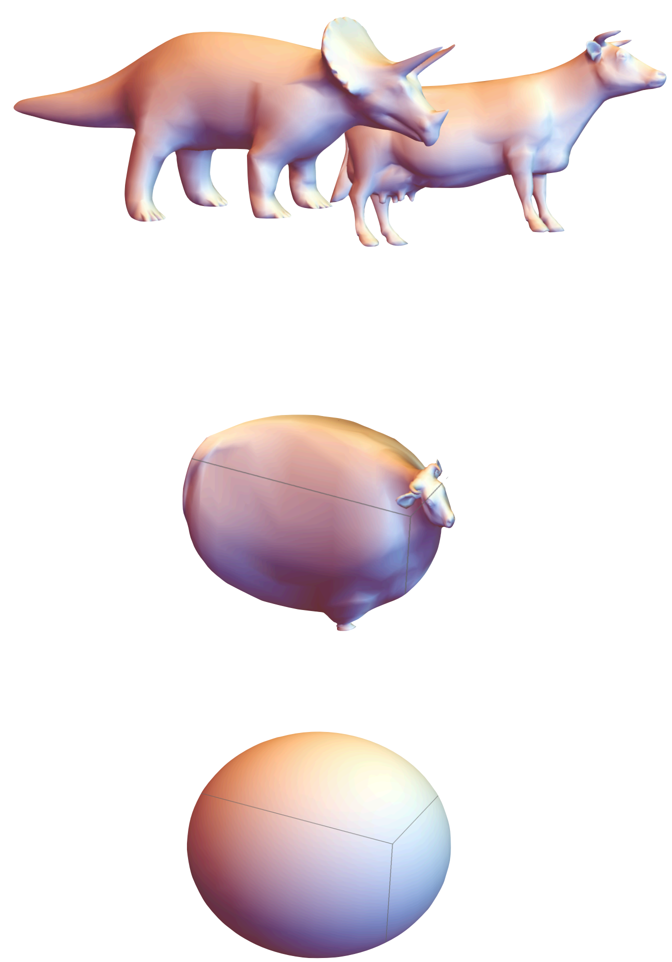


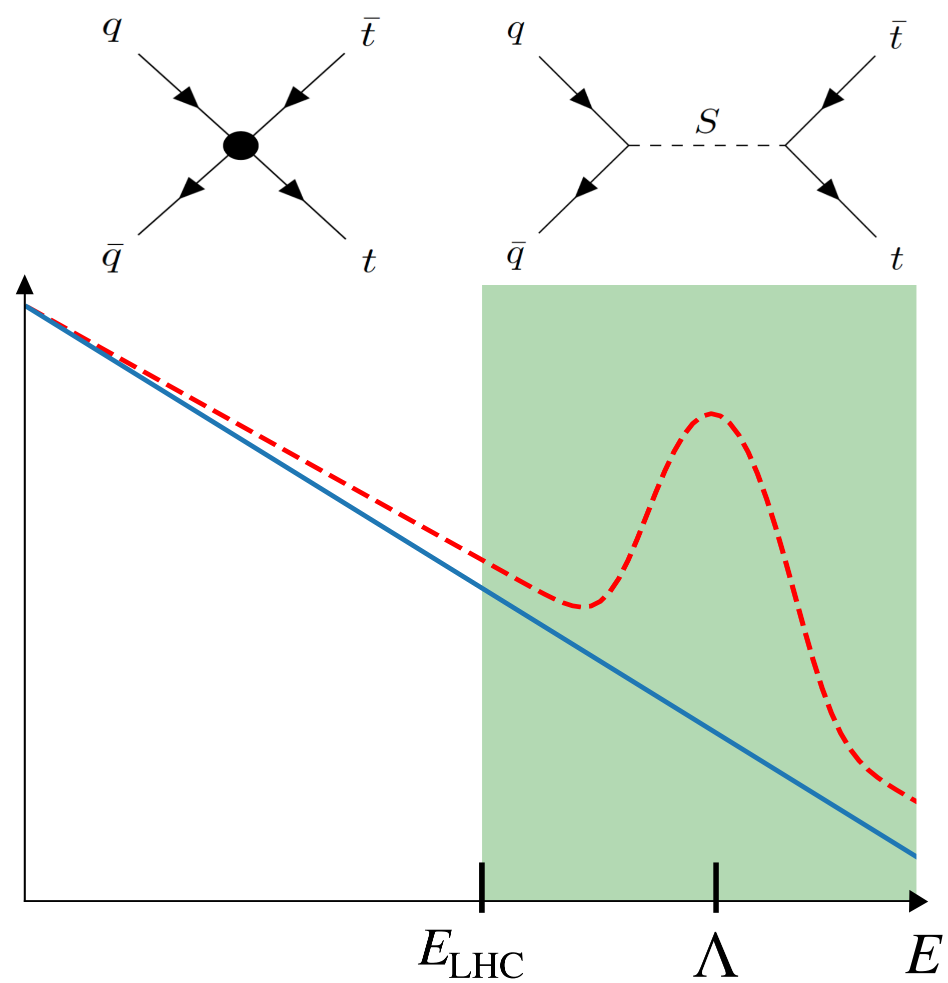


The Standard Model as an EFT

$$\mathscr{L}_{\rm SMEFT} = \mathscr{L}_{\rm SM} + \sum_i^{N_{d5}} \frac{c_i}{\Lambda}\mathcal{O}_i^{(5)} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_i^{N_{d7}} \frac{c_i}{\Lambda^3}\mathcal{O}_i^{(7)} + \sum_i^{N_{d8}} \frac{b_i}{\Lambda^4}\mathcal{O}_i^{(8)} + \ldots$$



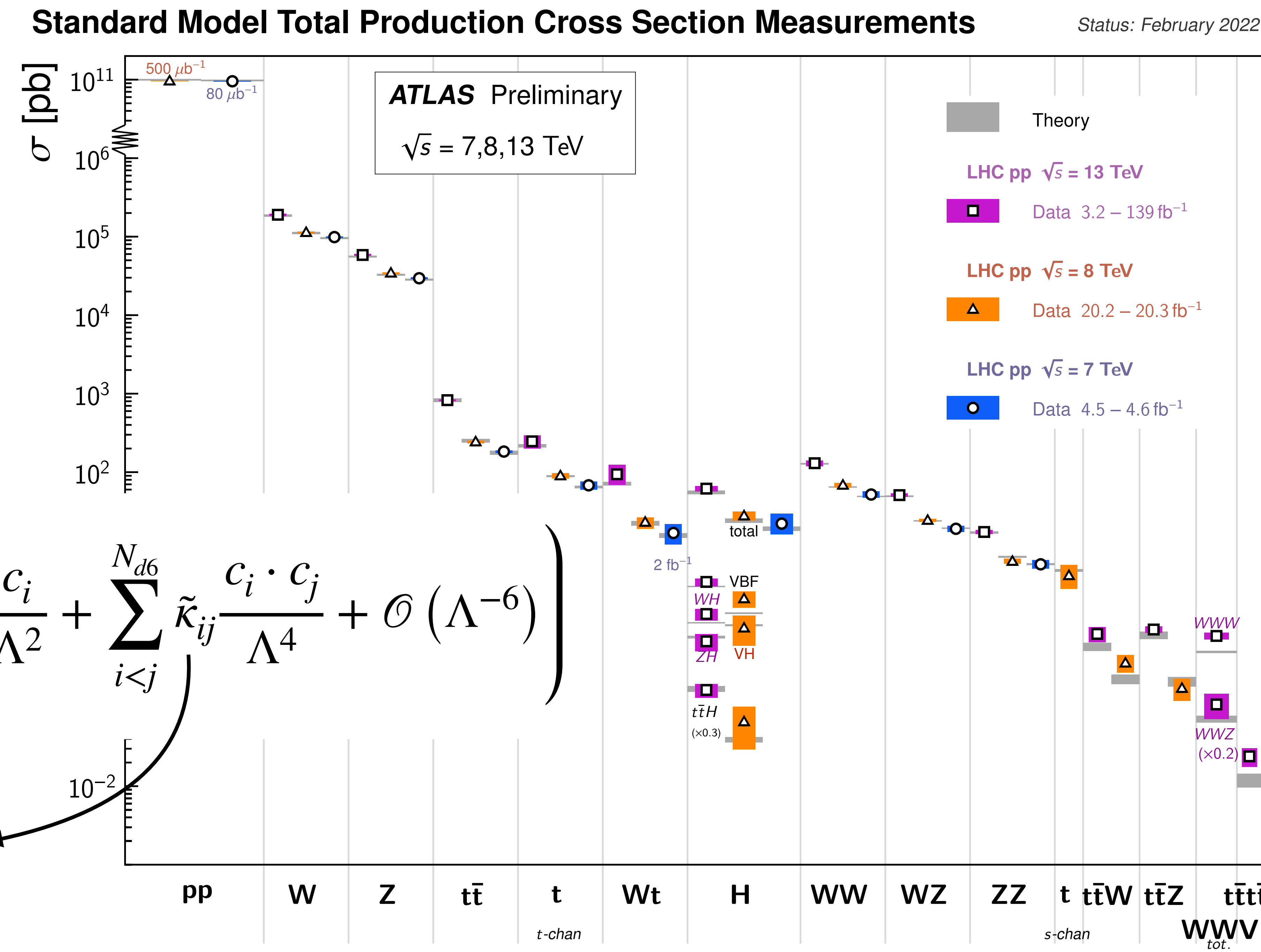




LHC observables in the SMEFT

Idea: parameterise (differential) cross-sections in terms of higher dimensional operators

[ATL-PHYS-PUB-2022-009]



Linear EFT corrections: interference SM-EFTd6 @NLO QCD

$$\sigma(c,\Lambda) = \sigma_{\rm SM} \times \left(1 + \sum_{i}^{N_{d6}} \kappa_i \frac{c_i}{\Lambda^2} + \sum_{i<j}^{N_{d6}} \tilde{\kappa}_{ij} \frac{c_i \cdot c_j}{\Lambda^4} + \mathcal{O}\left(\Lambda^{-6}\right)\right)$$

Quadratic EFT corrections: EFTd6-EFTd6 @NLO QCD

LHC observables in the SMEFT

From (differential) cross sections …

$$\sigma_{\mathrm{SMEFT}}(c, \Lambda) = \sigma_{\mathrm{SM}} \times \left(1 + \sum_{i}^{N_{d6}} \kappa_i \frac{c_i}{\Lambda^2} + \sum_{i<j}^{N_{d6}} \tilde{\kappa}_{ij} \frac{c_i \cdot c_j}{\Lambda^4} + \mathcal{O}\left(\Lambda^{-6}\right) \right)$$

Linear EFT corrections: interference SM-EFTd6 @NLO QCD

Quadratic EFT corrections: EFTd6-EFTd6 @NLO QCD

To a combined likelihood ready for optimisation …

$$-2 \log \mathscr{L} = \frac{1}{n_{\mathrm{dat}}} \sum_{i,j=1}^{n_{\mathrm{dat}}} \left(\sigma_{i,\mathrm{SMEFT}}(c) - \sigma_{i,\mathrm{exp}} \right) \left(\mathrm{cov}^{-1} \right)_{ij} \left(\sigma_{j,\mathrm{SMEFT}}(c) - \sigma_{j,\mathrm{exp}} \right)$$

Theory (pdf + scale) and experimental uncertainties (stat + systematics): $\mathrm{cov}^{(\mathrm{tot})}{}_{ij} = \mathrm{cov}^{(\mathrm{th})}{}_{ij} + \mathrm{cov}^{(\mathrm{exp})}{}_{ij}$

The SMEFiT framework

[2302.06660]

Theory

SM: **(N)NLO QCD + NLO EW**

EFT: **NLO QCD**, linear and quadratics, with **SMEFT@NLO**

NNPDF4.0 no top

Data

447 measurements from **Higgs, top, diboson** and **EWPO**

Full experimental **correlations**

Projections

HL-LHC, FCCee, CEPC

Automatic projection module

SMEFiT

[2309.04523]

UV models

External likelihoods

Fitting framework

Nested Sampling: Bayesian

Monte Carlo Replica Method (NNPDF)

External likelihoods

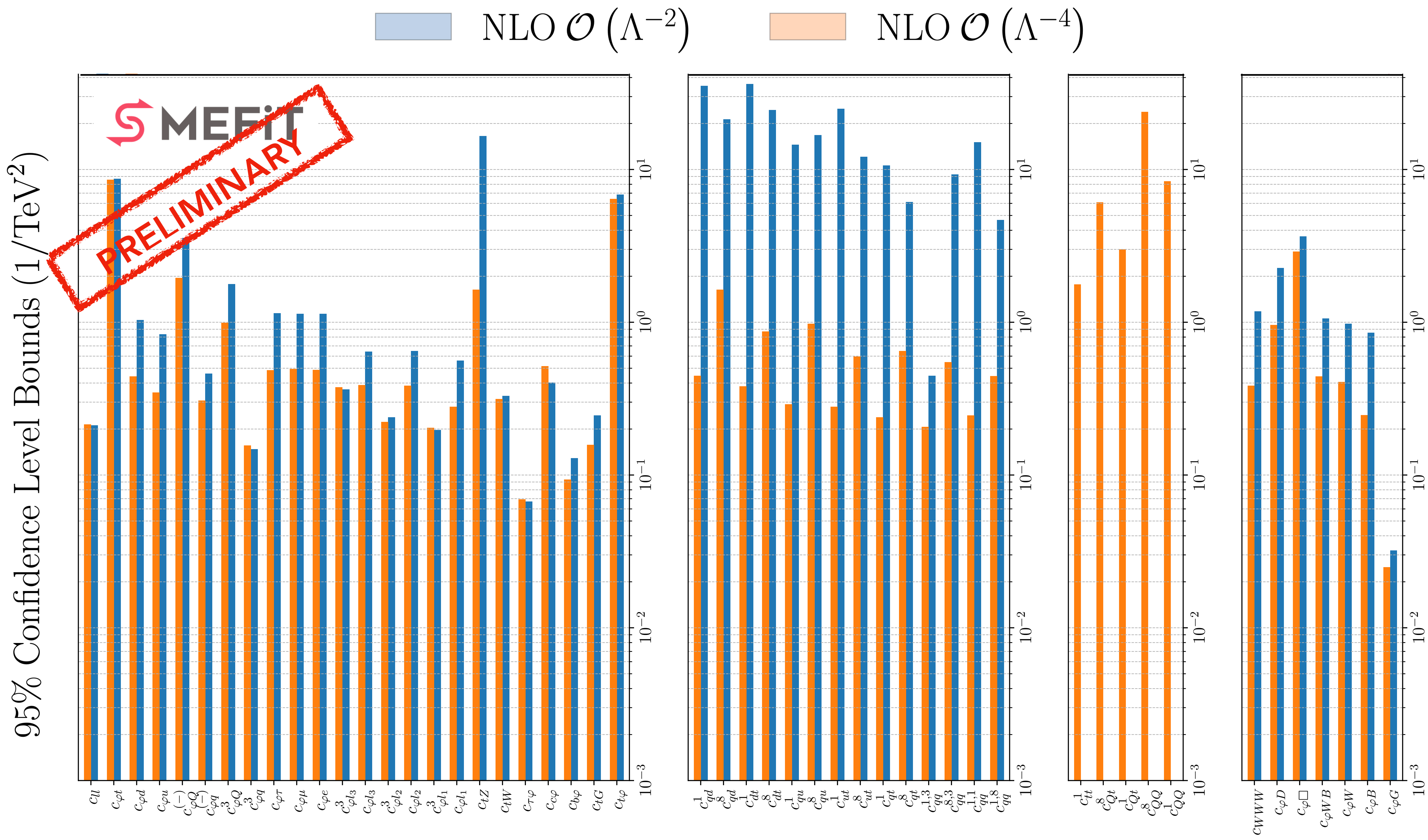
Validation

Extensive **toolbox**:

PCA, Fisher info, closure tests,..

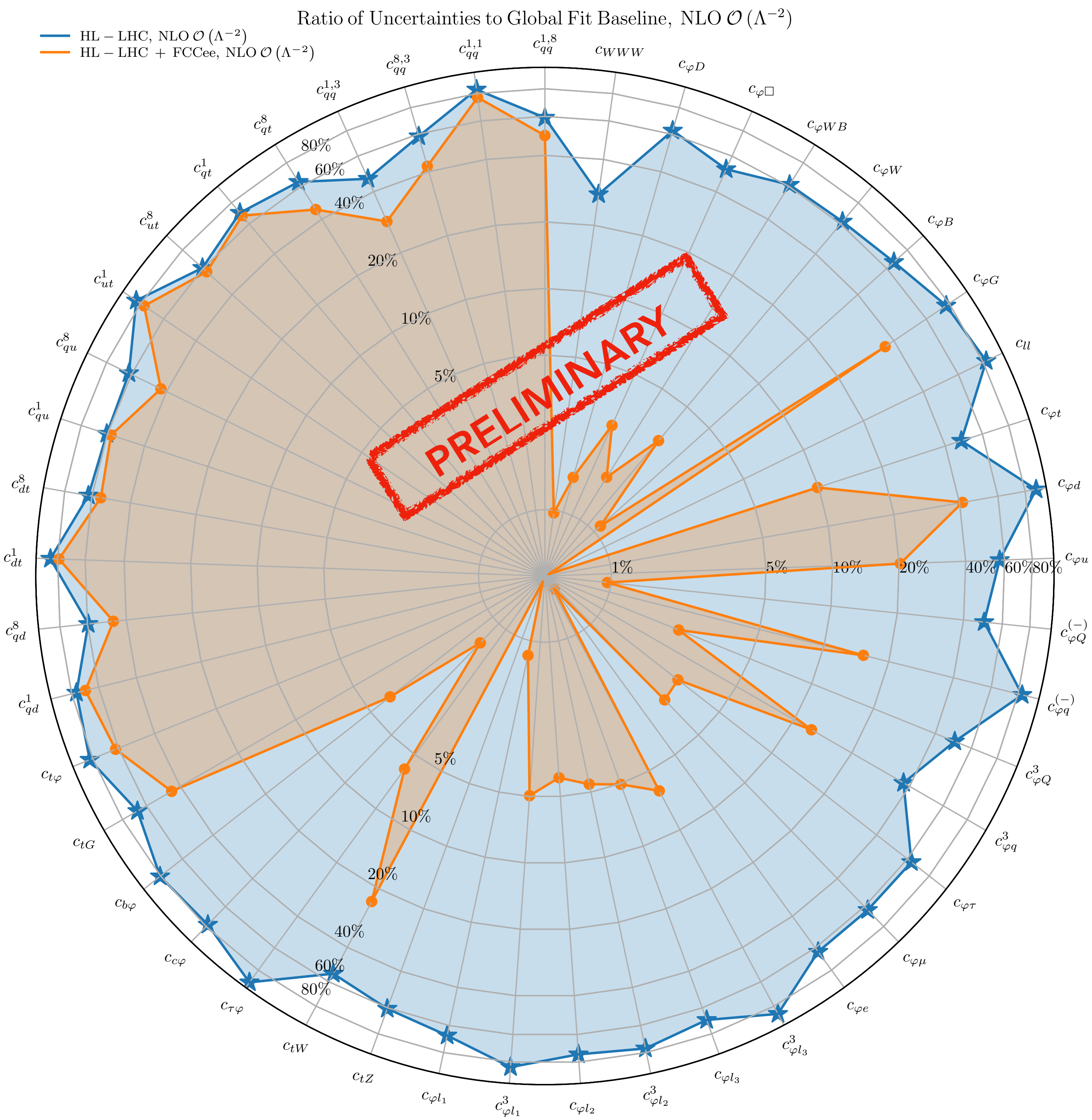
Full posterior distribution available

Result: state of the art



Global EFT fits include data on **top quark, Higgs, and gauge boson, EWPO,** both inclusive and differential measurements for a total of 447 measurements

Result: FCC projections



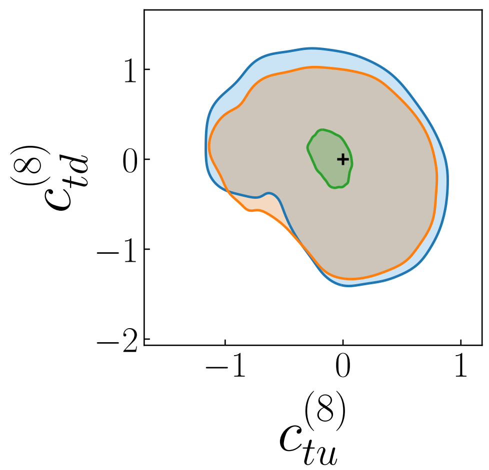
- The SMEFT will continue to be explored at future colliders
- Beyond the HL-LHC, a huge improvement is expected from the FCCee

But can we do (even) better?

- State of the art global efforts **reinterpret** "SM measurements" in an EFT context
- Which measurement is the most **sensitive** to EFT parameters?
  - Inclusive, single to multi-differential (which variables)
  - Binned or unbinned, which binning?

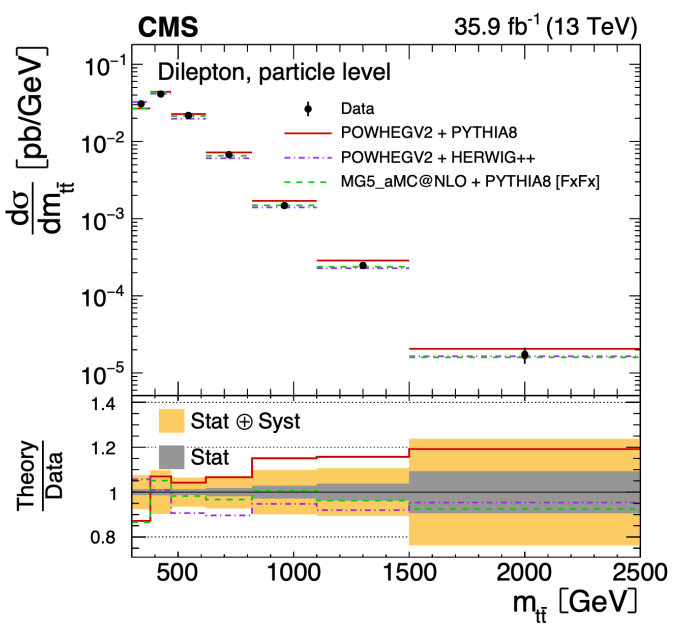
Framework needed to integrate **unbinned multivariate observables** into global SMEFT fits

- **Optimal** bounds on the EFT parameters
- Useful **diagnosis tool** to assess information loss

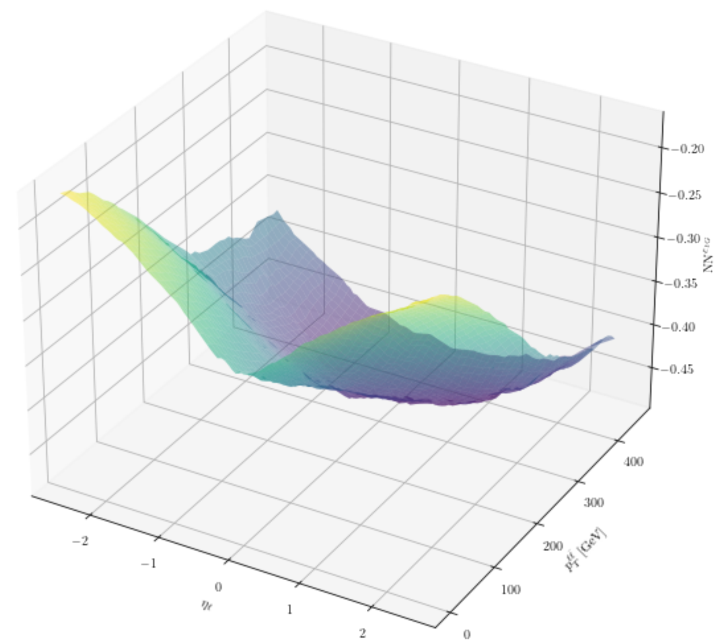


But can we do (even) better?

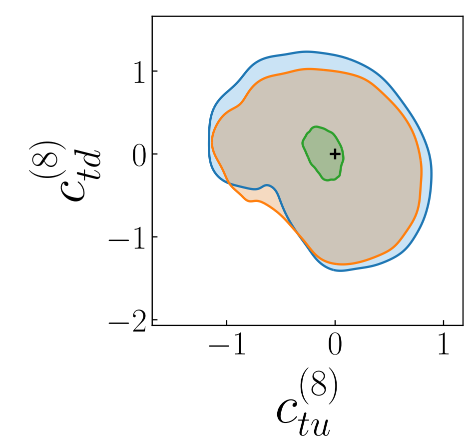
- State of the art global efforts **reinterpret** "SM measurements" in an EFT context



Binned, univariate

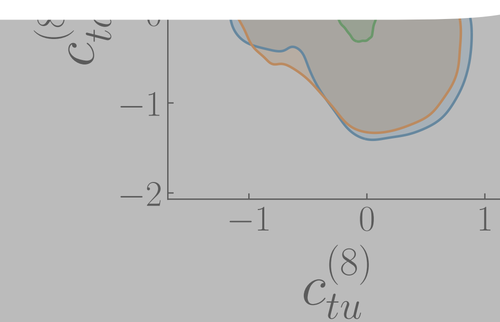


Unbinned, multivariate



Optimal inference

- **Optimal** bounds on the EFT parameters
- Useful **diagnosis tool** to assess information loss



In the remainder of this talk

✓ Mapping the SMEFT with SMEFiT

Unbinned multivariate observables: the ML4EFT framework

Case study: ML observables in the Higgs and top sector

A combined ML4EFT + SMEFiT analysis

ML4EFT

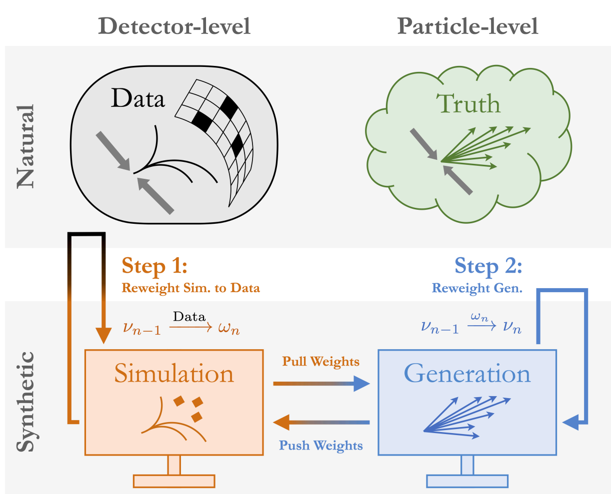
[2211.02058] R. Gomez Ambrosio, JtH, M. Madigan, J. Rojo, V.Sanz

https://lhcfitnikhef.github.io/ML4EFT

Where theory and experiment meet

We are progressively moving through the simulation chain (latent space)

$$p(x|c) \sim \int dz_{\text{det}} dz_{\text{shower}} dz_{\text{parton}} p(x|z_{\text{det}}) p(z_{\text{det}}|z_{\text{shower}}) p(z_{\text{shower}}|z_{\text{parton}}) p(z_{\text{parton}}|c)$$



DELPHES
fast simulation

SMEFT@NLO

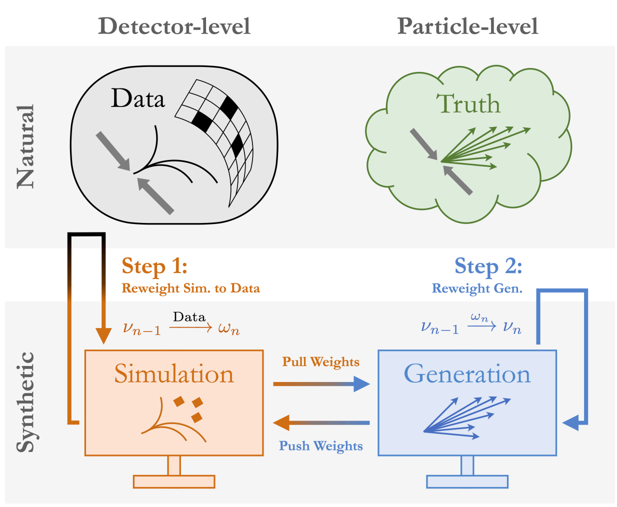
Unbinned unfolding, Omnifold [1911.09107]

✓	**Generation**	Observed	←	Hidden/latent variables	←	POI
?	**(Likelihood free) Inference**	Observed	→	Hidden/latent variables	→	POI

Where theory and experiment meet

We are progressively moving through the simulation chain (latent space)

$$p(x|c) \sim \int dz_{\rm det} dz_{\rm shower} dz_{\rm parton} p(x|z_{\rm det}) p(z_{\rm det}|z_{\rm shower}) p(z_{\rm shower}|z_{\rm parton}) p(z_{\rm parton}|c)$$



DELPHES fast simulation

SMEFT@NLO

Unbinned unfolding, Omnifold [1911.09107]

Generation Observed ← Hidden/latent variables ← POI

? **(Likelihood free) Inference** Observed → Hidden/latent variables → POI

Likelihood free inference

- Starting from two balanced datasets $\mathscr{D}_{\mathrm{SM}}$ and $\mathscr{D}_{\mathrm{EFT}}$ drawn from $f(\boldsymbol{x}\,|\,\mathrm{SM})$ and $f(\boldsymbol{x}\,|\,\mathrm{EFT})$, we minimise e.g. the cross-entropy loss

$$L[g(\boldsymbol{x})] = -\frac{1}{N}\sum_{e\in\mathscr{D}_{\mathrm{EFT}}} w_e \log(1-g(\boldsymbol{x}_e)) - \frac{1}{N}\sum_{\mathscr{D}_{\mathrm{SM}}} w_e \log g(\boldsymbol{x}_e)$$

Event weights

$\{m_{t\bar{t}}, \eta_l, \Delta\phi, \ldots\}$

- The learned decision boundary $g(\boldsymbol{x})$ is one-to-one with the likelihood ratio (LR) as $N \to \infty$

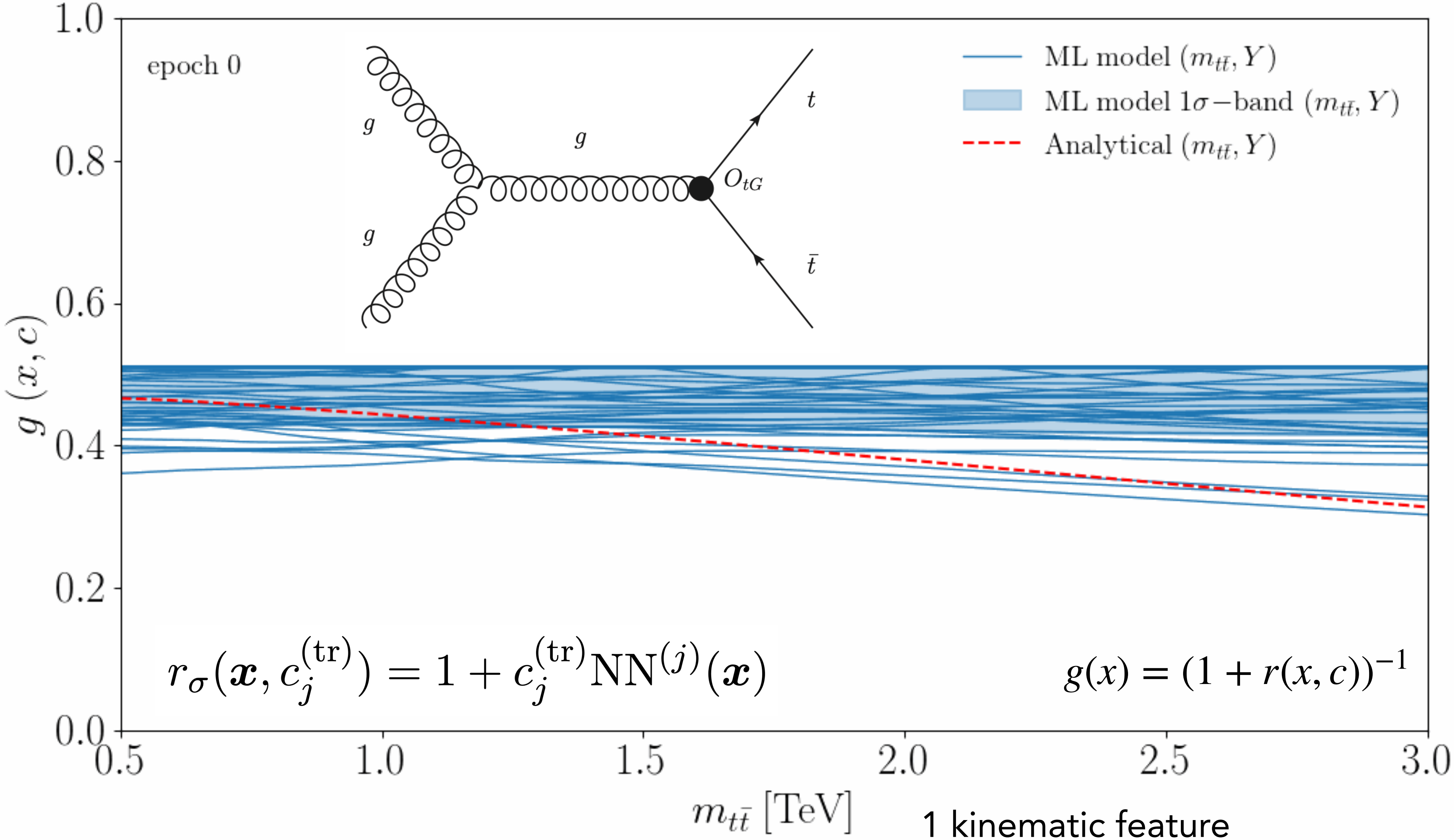
$$\frac{\delta L}{\delta g} = 0 \implies \hat{g}(\boldsymbol{x}) = \left(1 + \frac{f(\boldsymbol{x}\,|\,\mathrm{EFT})}{f(\boldsymbol{x}\,|\,\mathrm{SM})}\right)^{-1} \equiv \frac{1}{1+r(\boldsymbol{x})}$$

Parameterise with NNs

- Parameterising $g(\mathbf{x})$ inside L with NNs lets us extract the likelihood (ratio) **implicitly**

Likelihood learning in practice

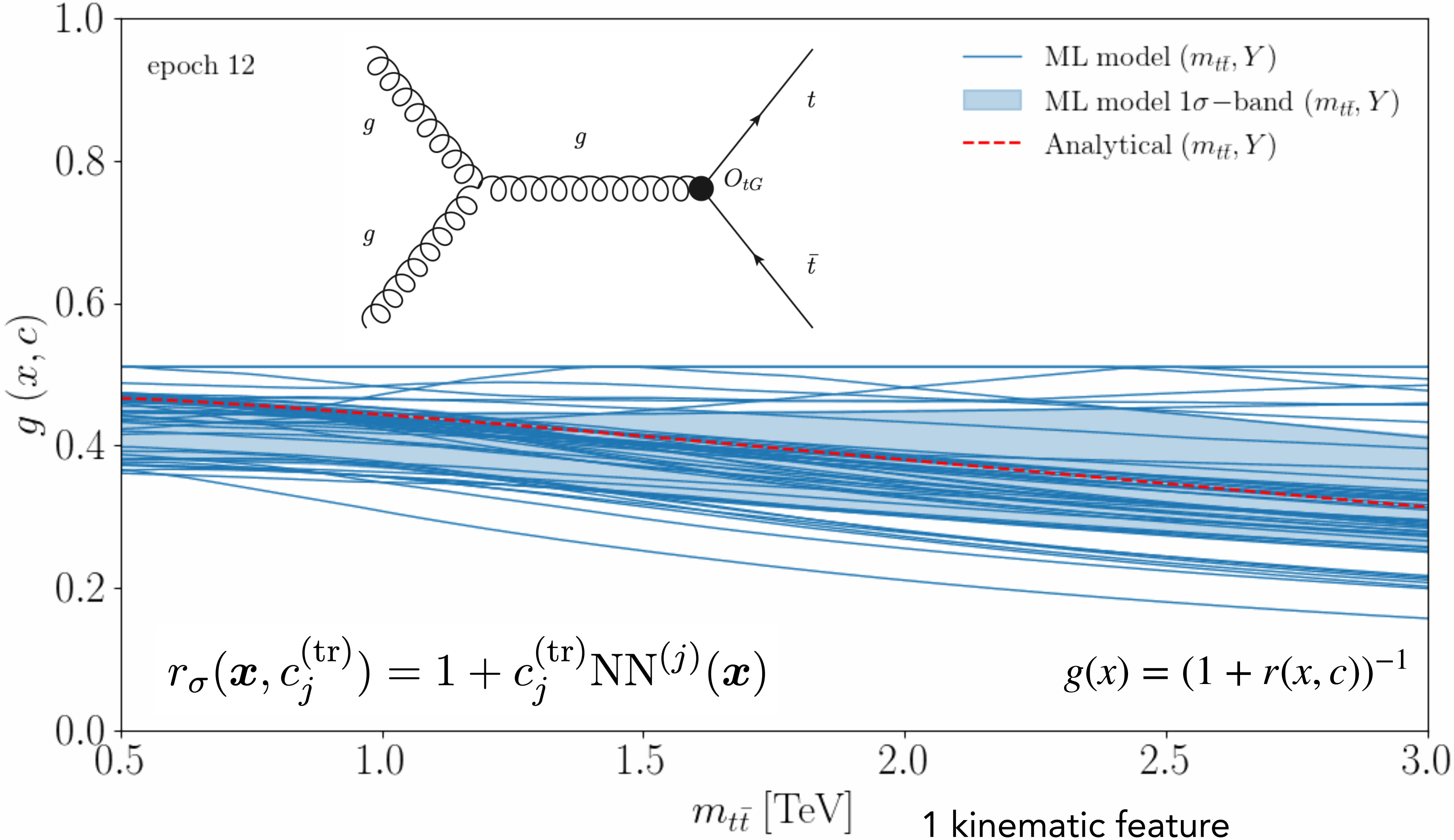
$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\rm eft}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\rm sm}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$



1 kinematic feature

Likelihood learning in practice

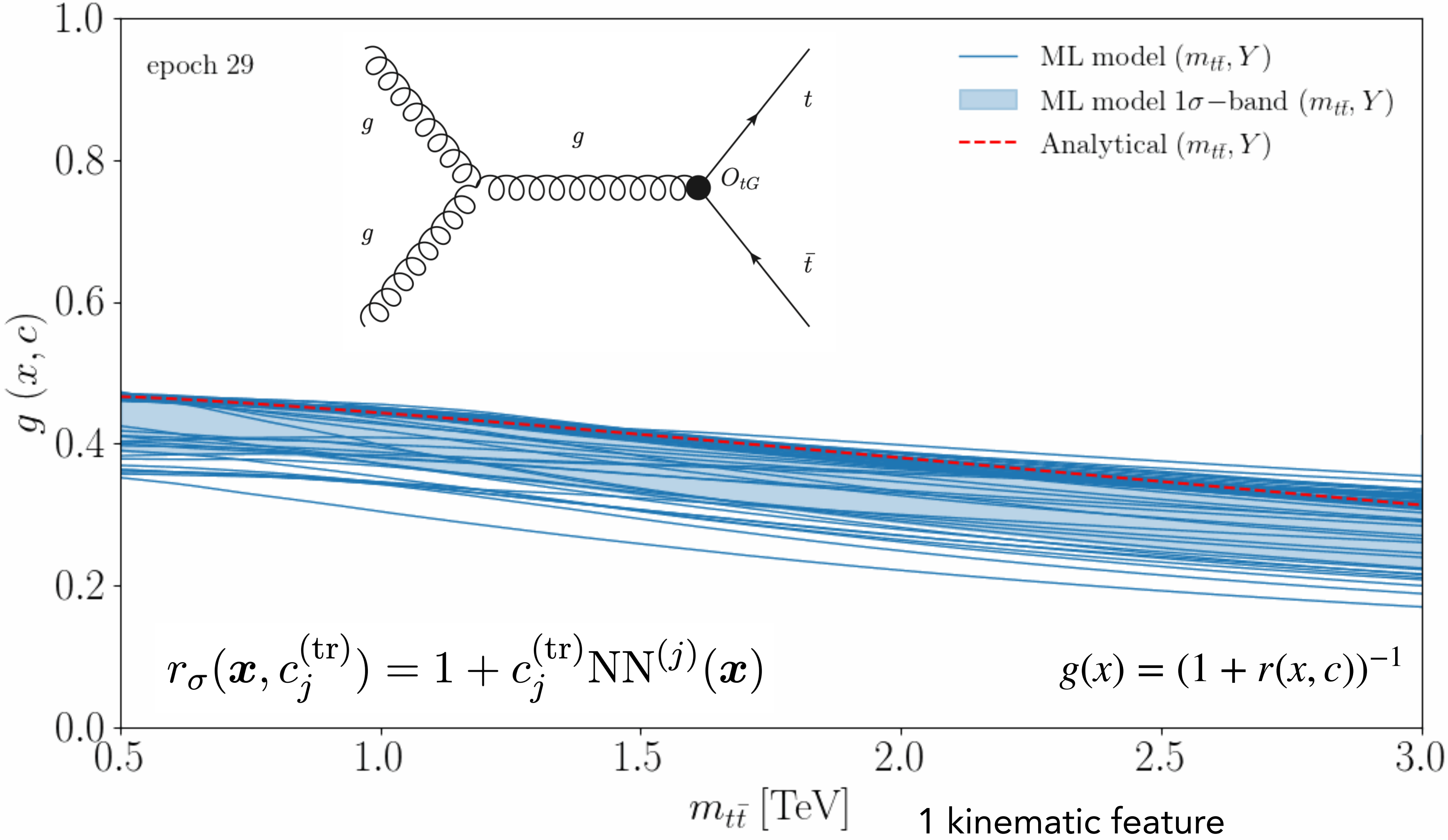
$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\text{eft}}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\text{sm}}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$



1 kinematic feature

Likelihood learning in practice

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\rm eft}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\rm sm}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$



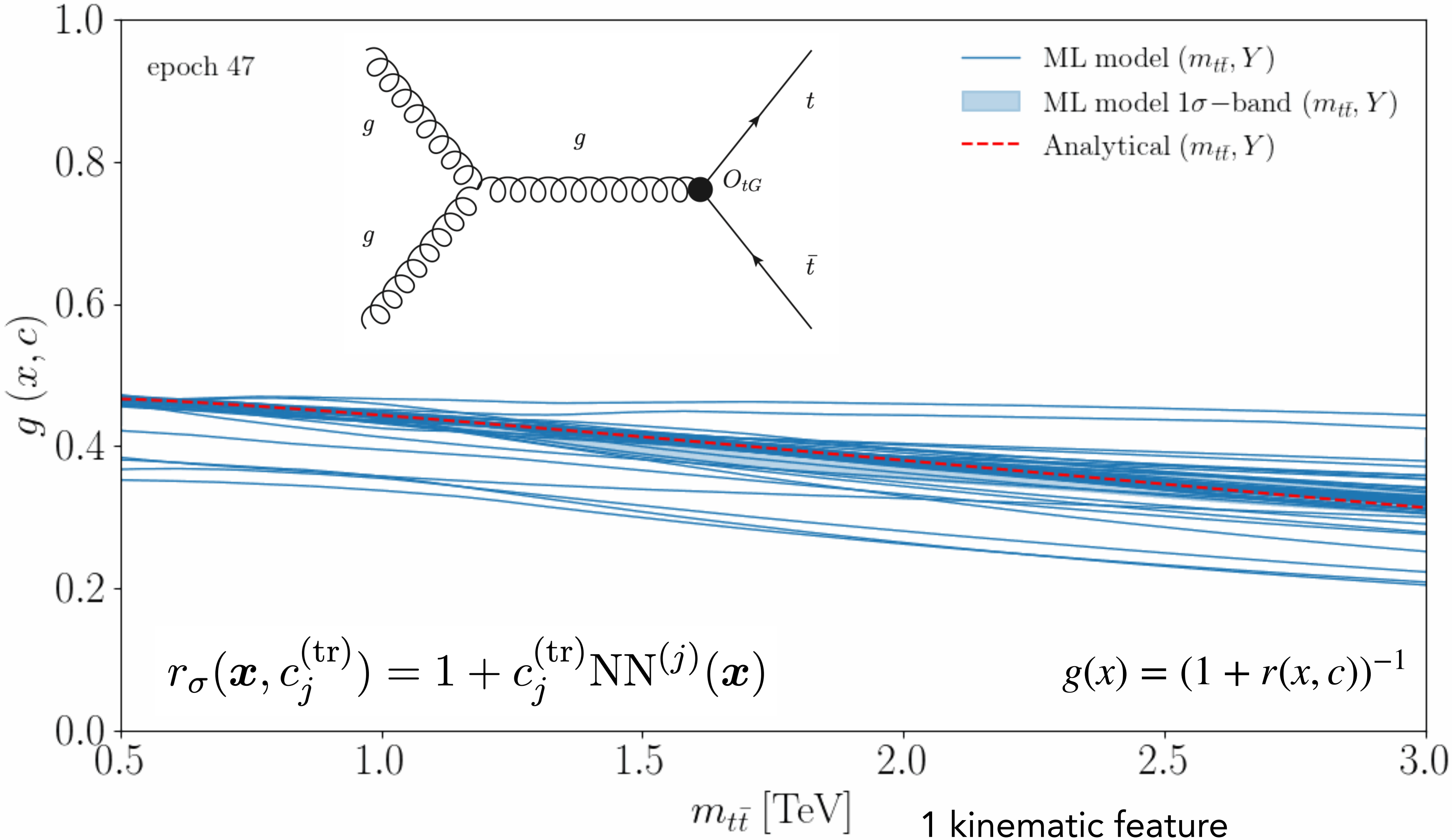
$$r_\sigma(\boldsymbol{x}, c_j^{(\rm tr)}) = 1 + c_j^{(\rm tr)}\mathrm{NN}^{(j)}(\boldsymbol{x})$$

$$g(x) = (1 + r(x,c))^{-1}$$

1 kinematic feature

Likelihood learning in practice

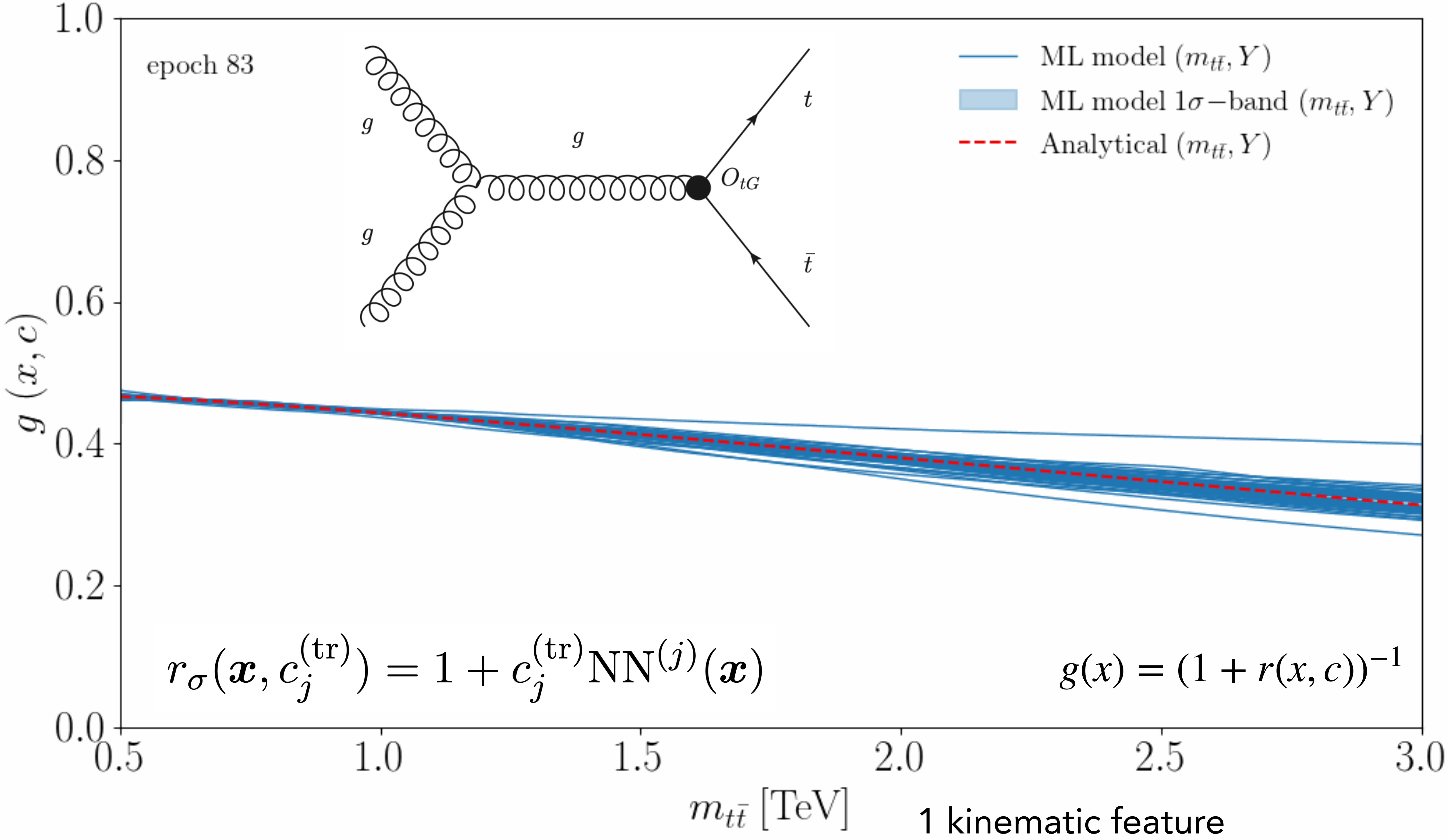
$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\rm eft}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\rm sm}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$



1 kinematic feature

Likelihood learning in practice

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\rm eft}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\rm sm}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$



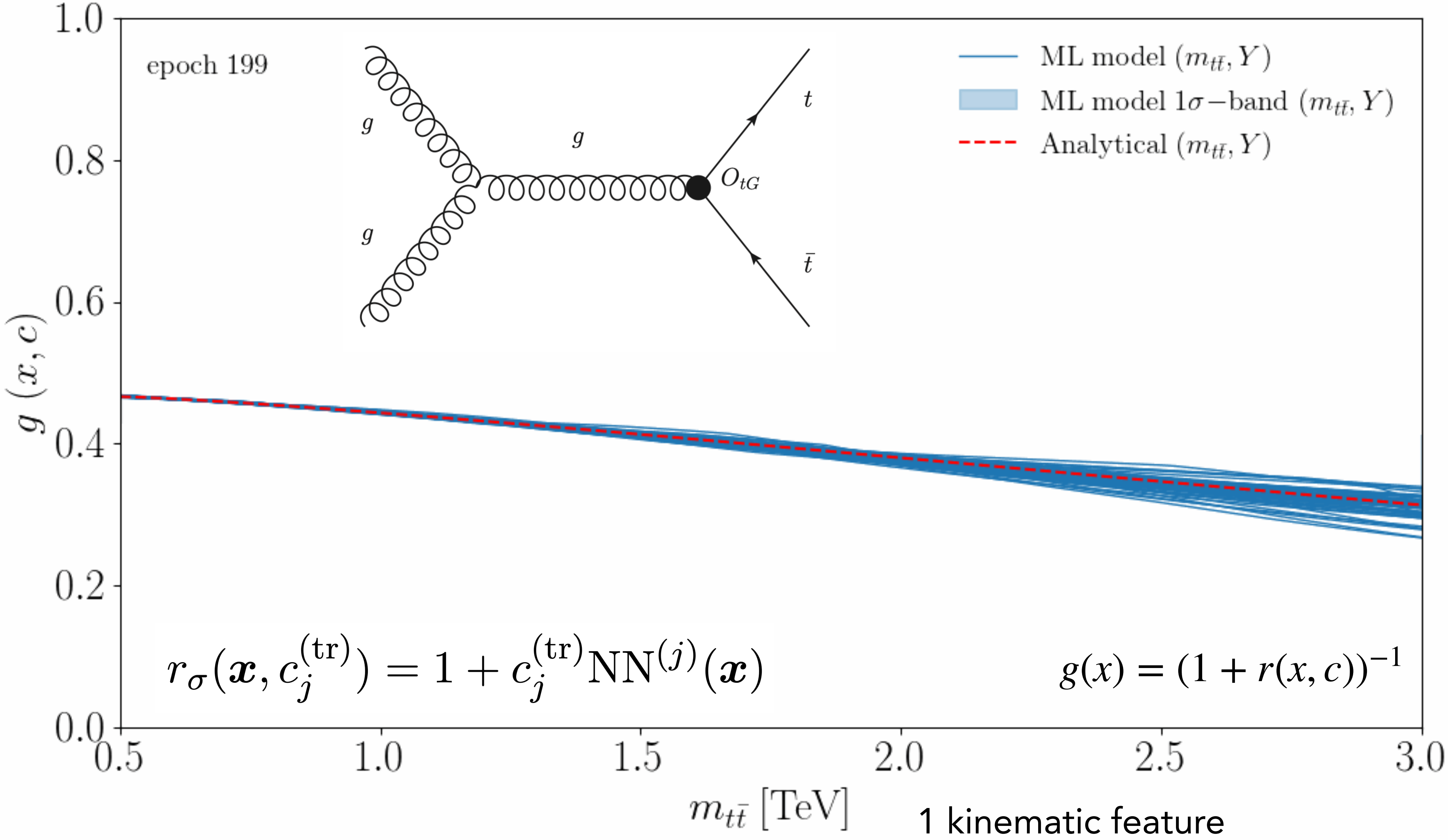
$$r_\sigma(\boldsymbol{x}, c_j^{(\rm tr)}) = 1 + c_j^{(\rm tr)}\mathrm{NN}^{(j)}(\boldsymbol{x})$$

$$g(x) = (1 + r(x,c))^{-1}$$

1 kinematic feature

Likelihood learning in practice

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\rm eft}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\rm sm}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$

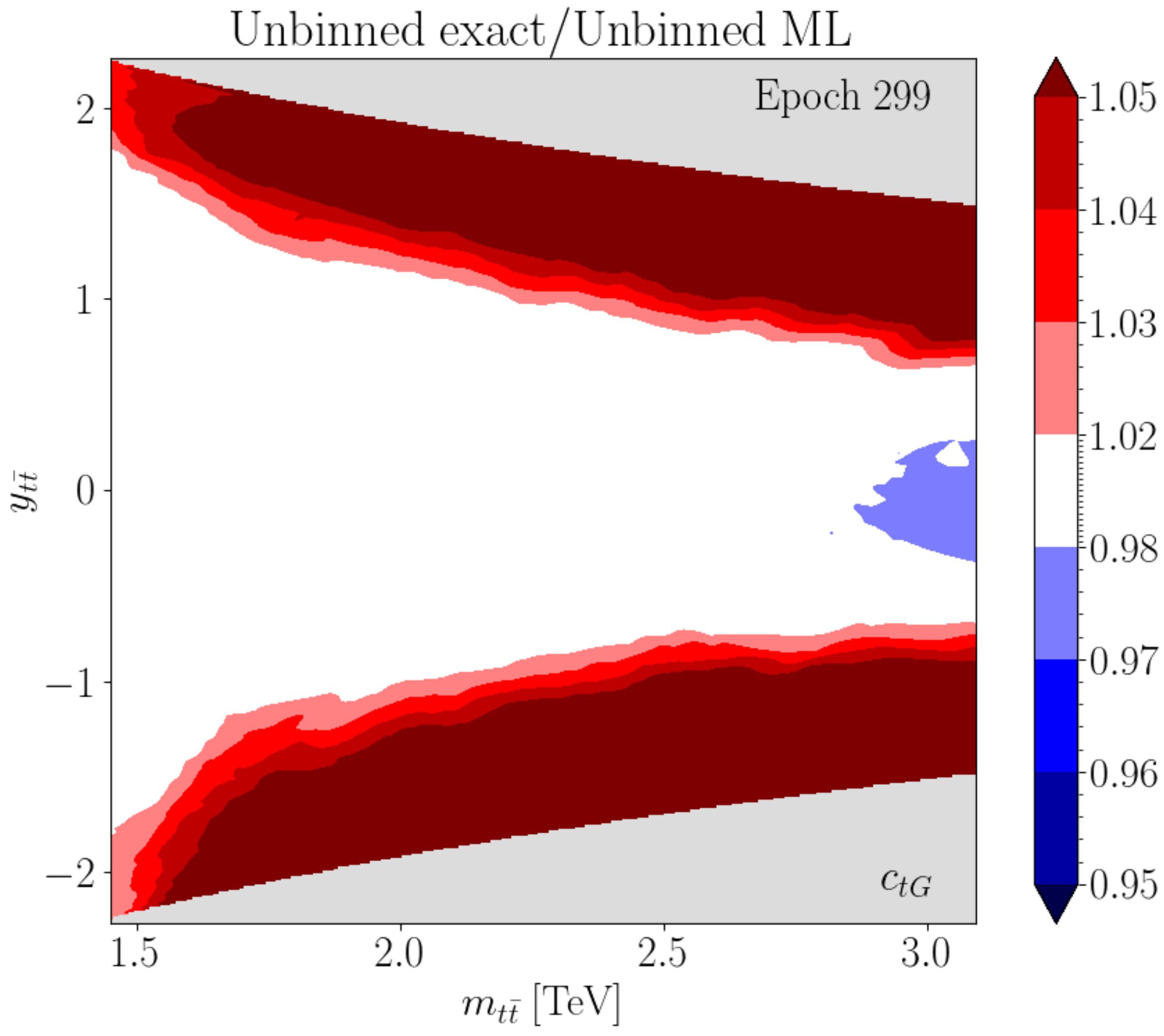


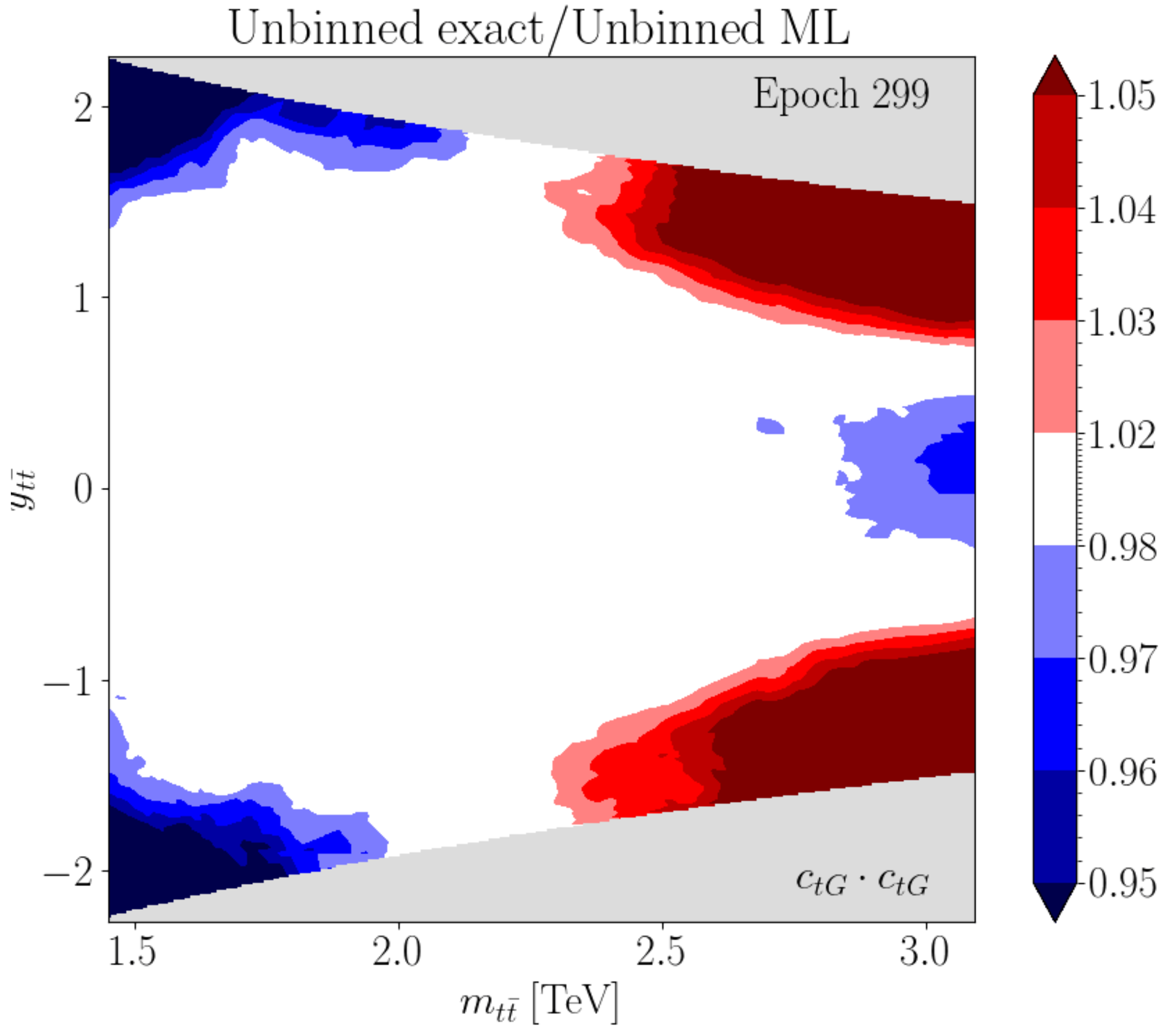
1 kinematic feature

Likelihood learning in practice

2 kinematic features

$$r_\sigma(\boldsymbol{x}, c_j^{(\mathrm{tr})}) = 1 + c_j^{(\mathrm{tr})}\mathrm{NN}^{(j)}(\boldsymbol{x})$$





The ML4EFT framework

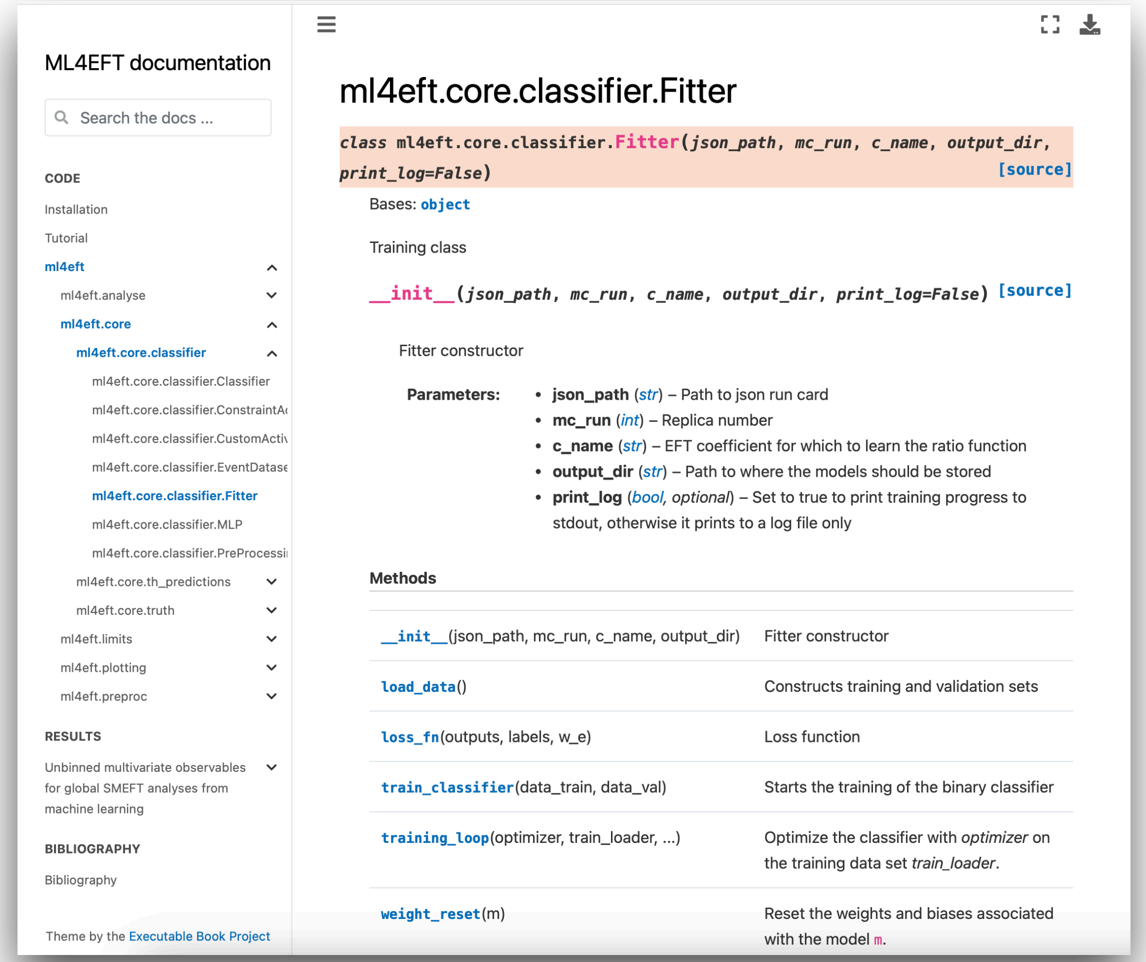
`pip install ml4eft`

https://lhcfitnikhef.github.io/ML4EFT

2211.02058 R. Gomez Ambrosio, JtH, M. Madigan, J. Rojo, V.Sanz

Open-source NN-based python framework for the integration of unbinned multivariate observables into global SMEFT fits

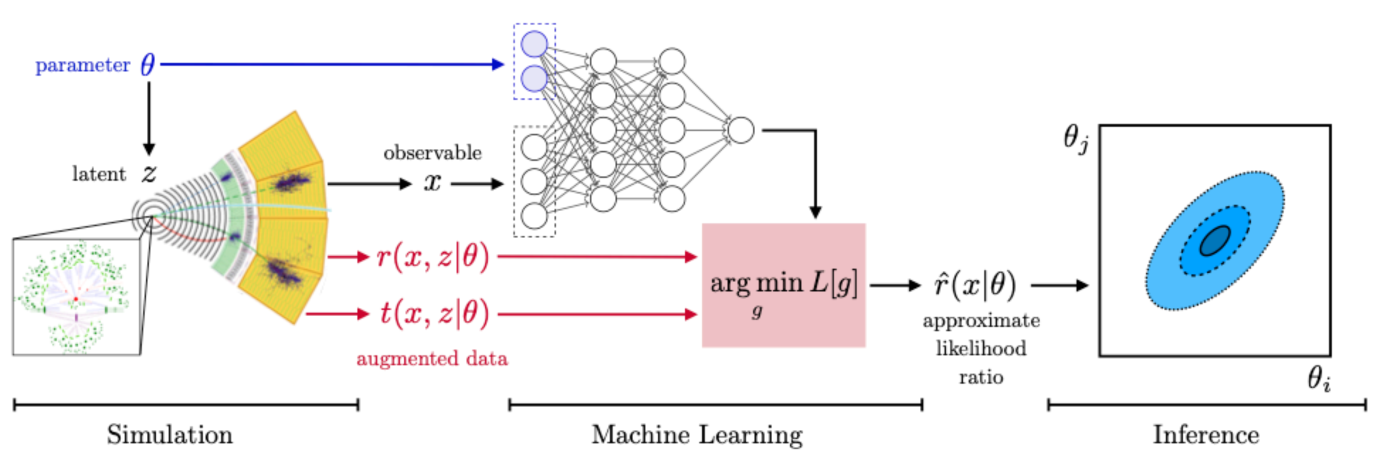
- **Goal**: provide optimal constraints on the SMEFT
- **Diagnostic tool:** what is the information loss incurred by a particular choice of bins?
- **Projections:** how will SMEFT constraints improve if unbinned data are made available?



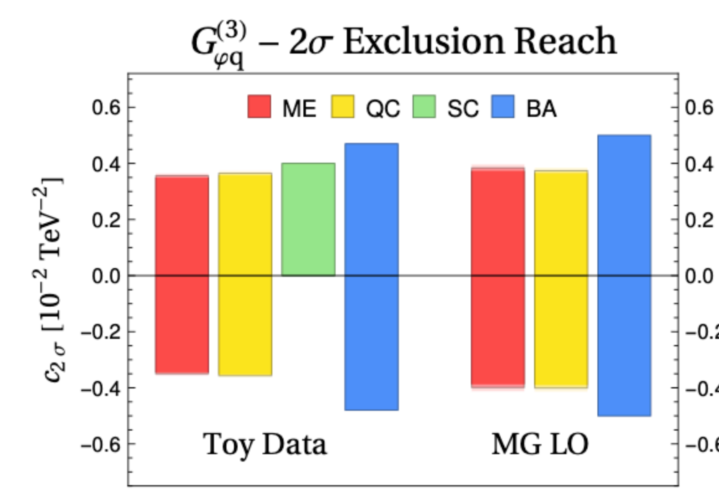
Modular structure, easy to maintain, well documented

Related work

- **Madminer series** (J.Brehmer, K.Cranmer, G.Louppe et al.) [1907.10621, 1805.00020, …]
- **Parameterized classifiers for SMEFT** (A. Glioti et al.) [2007.10356]
- **Learning the EFT likelihood with tree boosting** (R. Schöfbeck et al) [2205.12976]
- **Back to the Formula** (A. Butter, T. Plehn et al) [2109.10414]
- **Boosted likelihood learning with event reweighing** (A. Glioti et al) [2308.05704]
- **Designing Observables for Measurements with Deep Learning** (O.Long, B. Nachman) [2310.08717]



[2010.06439]



[2007.10356]

Anticipating global fits

- Global EFT fits typically feature ~50 WCs and thus efficient scaling with the number of WCs becomes essential

- Solution: learn the coefficient functions separately and combine afterwards

$$r(\boldsymbol{x}, \boldsymbol{c}) \equiv \frac{d\sigma(\boldsymbol{x}, \boldsymbol{c})}{d\sigma(\boldsymbol{x}, \boldsymbol{0})} = 1 + \sum_{j=1}^{n_{\rm eft}} r^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k \geq j}^{n_{\rm eft}} r^{(j,k)}(\boldsymbol{x}) c_j c_k$$

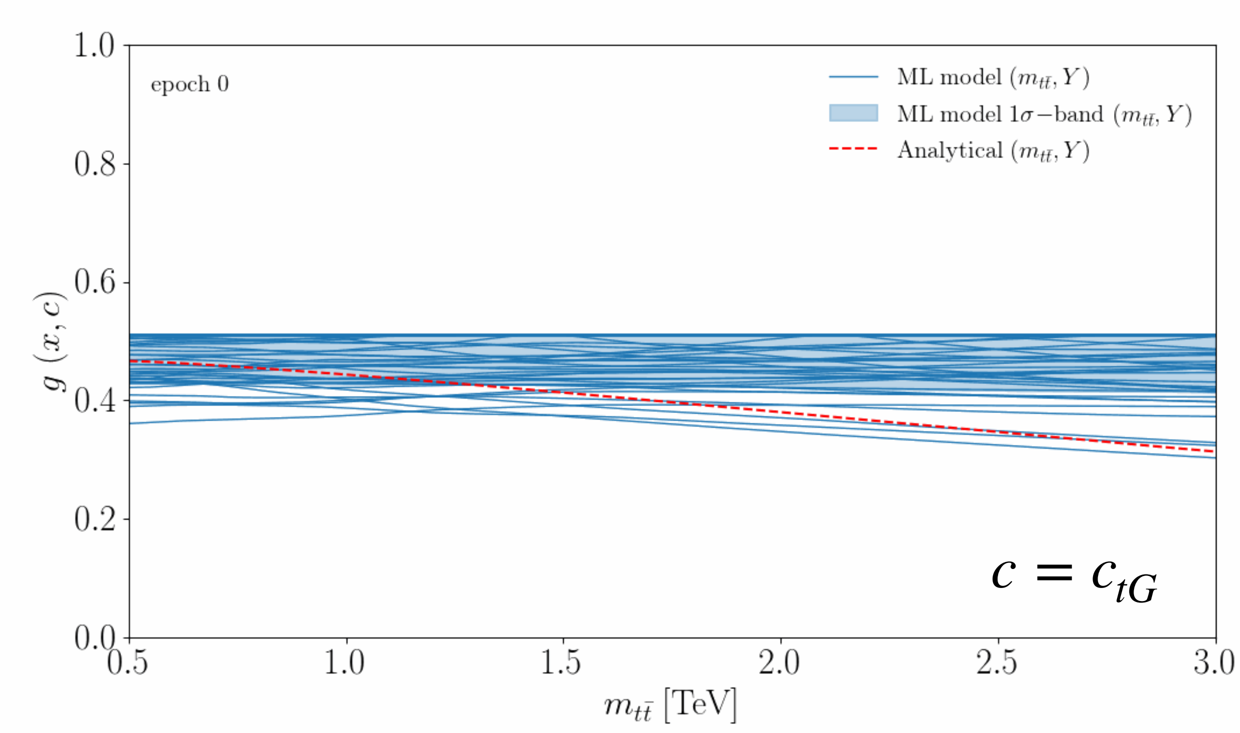
Example: to learn a single $r^{(j)}$, generate $\mathscr{D}_{\rm sm}$ and $\mathscr{D}_{\rm eft}$ at c_j up to $\mathcal{O}(\Lambda^{-2})$. Then $r(\boldsymbol{x}, \boldsymbol{c}) = 1 + r^{(j)}(\boldsymbol{x}) c_j^{(\rm tr)}$ and training means

$$g(\boldsymbol{x}, c_j^{(\rm tr)}) = \left(1 + \left[1 + c_j^{(\rm tr)} \cdot \mathrm{NN}^{(j)}(\boldsymbol{x})\right]\right)^{-1} \qquad \mathrm{NN}^{(j)}(\boldsymbol{x}) \to r^{(j)}(\boldsymbol{x})$$

Uncertainty treatment

- Stick to a regime in which statistical uncertainties dominate over systematics
- Finite training data makes one subject to methodological uncertainties
- Solution: propagate uncertainties to the space of models by training multiple replicas

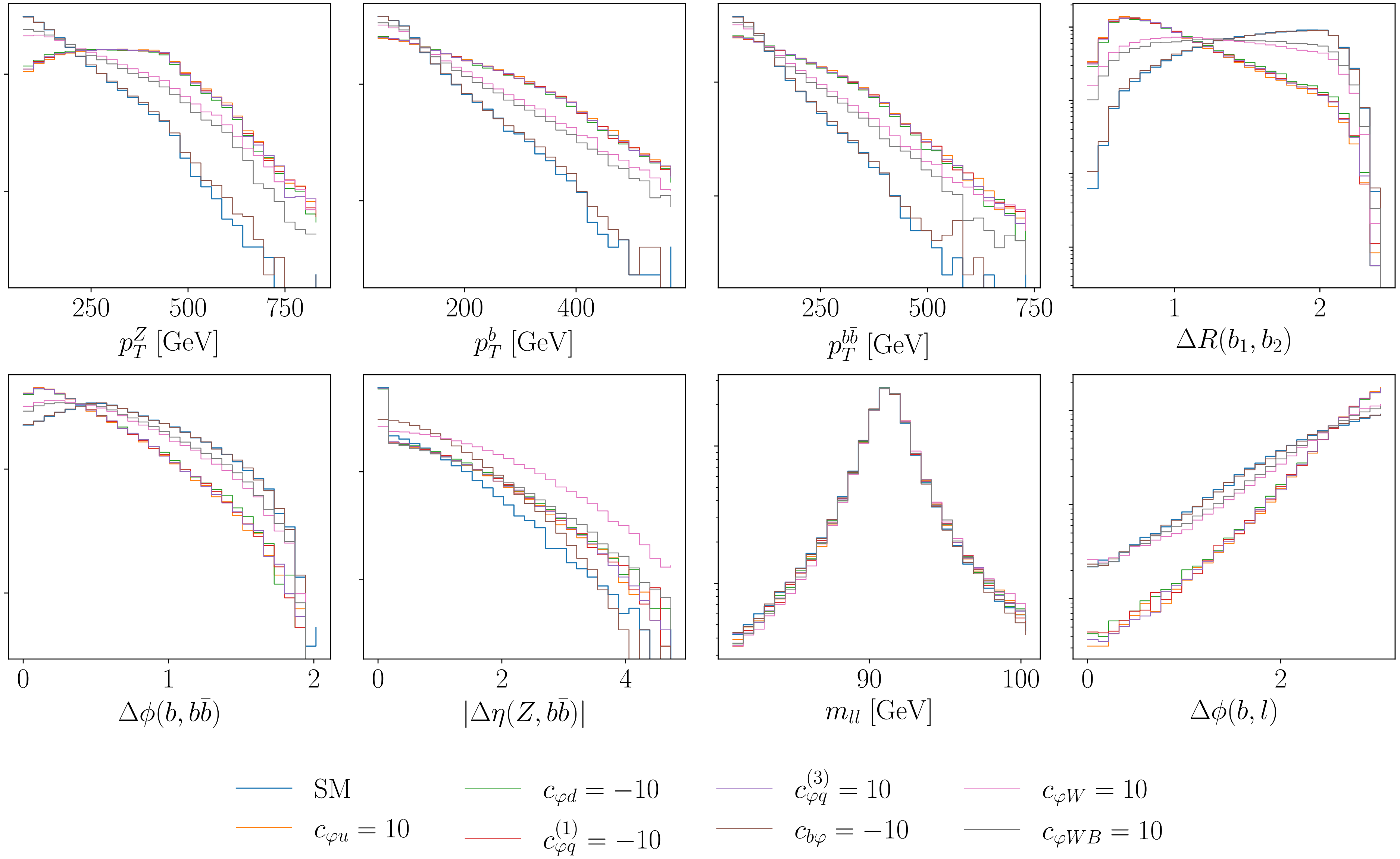
$$\hat{r}^{(i)}(\boldsymbol{x}, \boldsymbol{c}) \equiv 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}_i^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k \geq j}^{n_{\rm eft}} \mathrm{NN}_i^{(j,k)}(\boldsymbol{x}) c_j c_k \,, \qquad i = 1, \ldots, N_{\rm rep}$$



Process	$N_{\rm rep}$	$\widetilde{N}_{\rm ev}$ (per replica)	$N_{\rm nn}$	#trainings
$pp \to t\bar{t}$	50	10^5	4	200
$pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$	25	10^5	40	1000

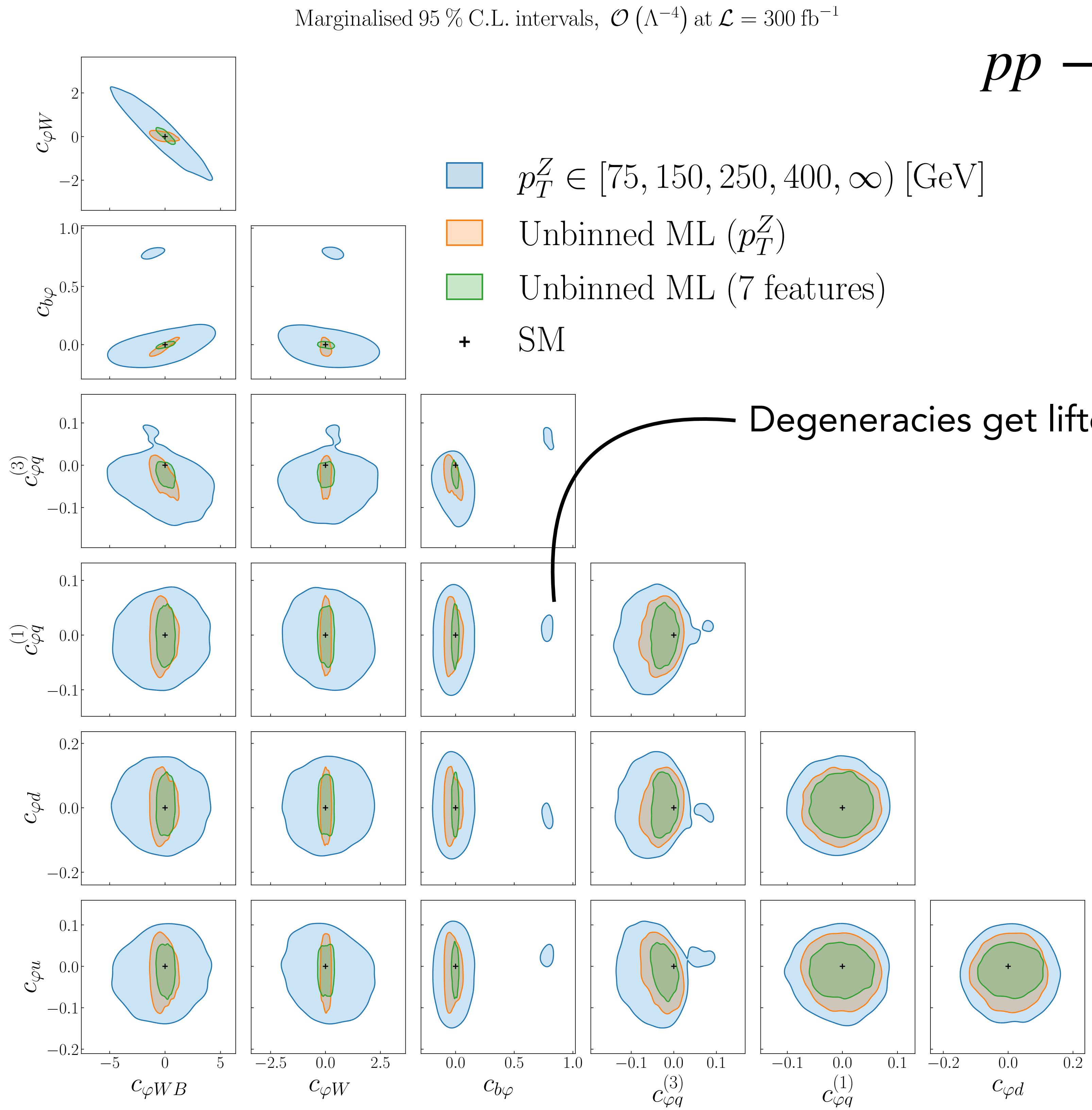
Let's go multivariate

- $pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$: 18 features, 8 EFT coefficients
- $pp \to hZ \to b\bar{b}\ell^+\ell^-$: 7 features, 7 EFT coefficients



Results: Higgs + Z associated production

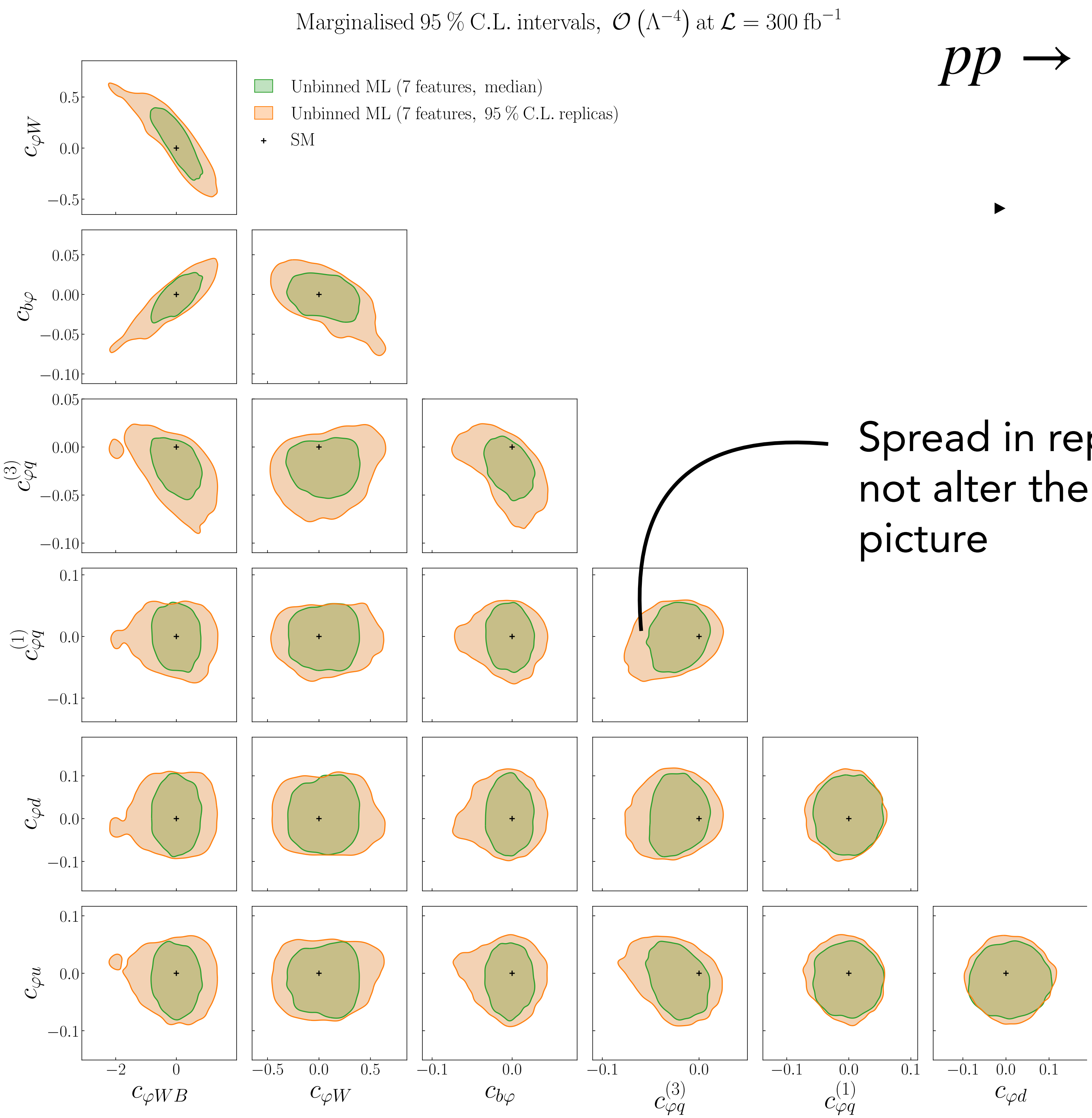
$$pp \to hZ \to b\bar{b}\ell^+\ell^-$$



Degeneracies get lifted!

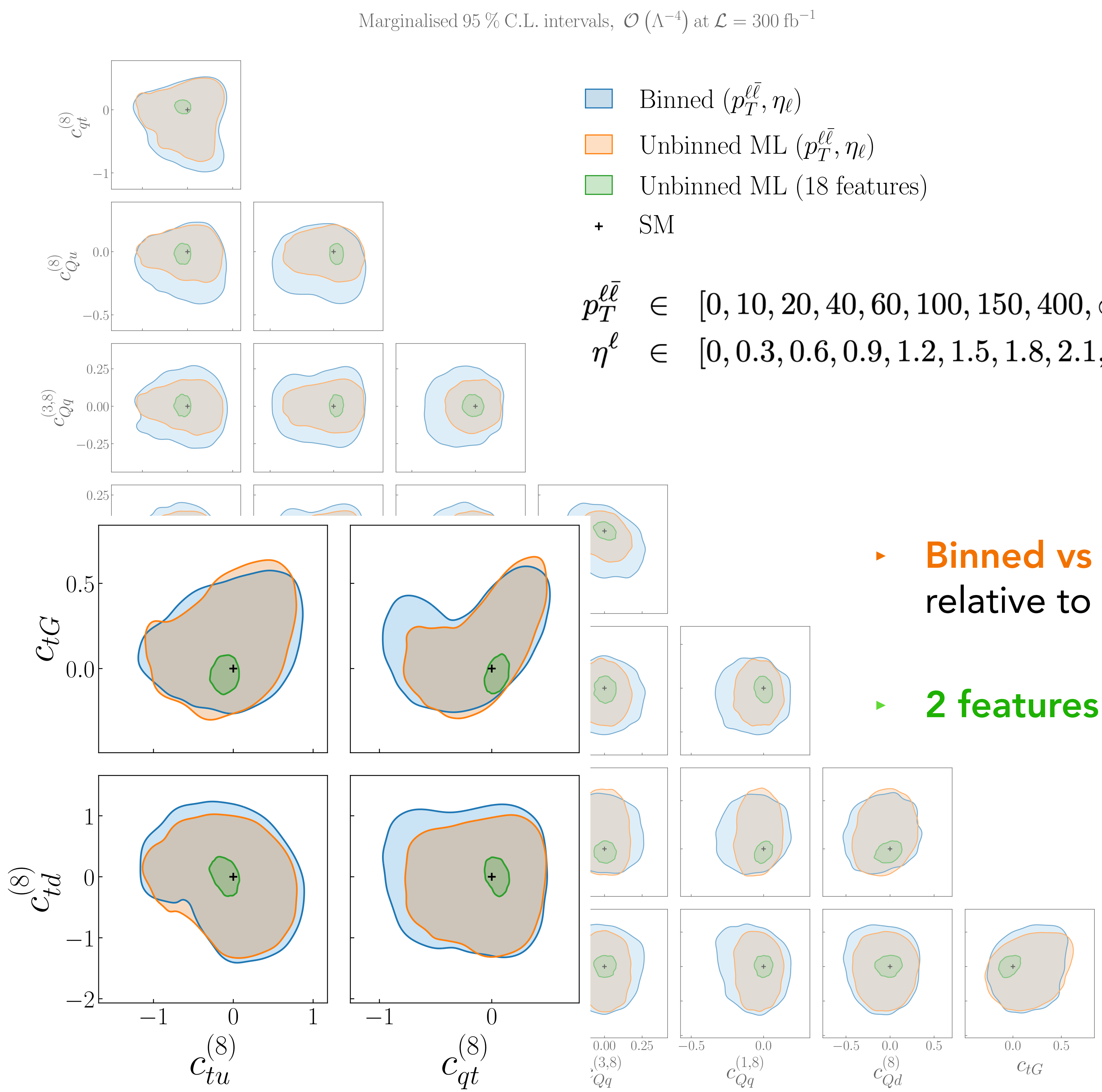
Results: Higgs + Z associated production

$$pp \to hZ \to b\bar{b}\ell^+\ell^-$$



Spread in replicas does not alter the qualitative picture

Results: unbinned observables in the top sector



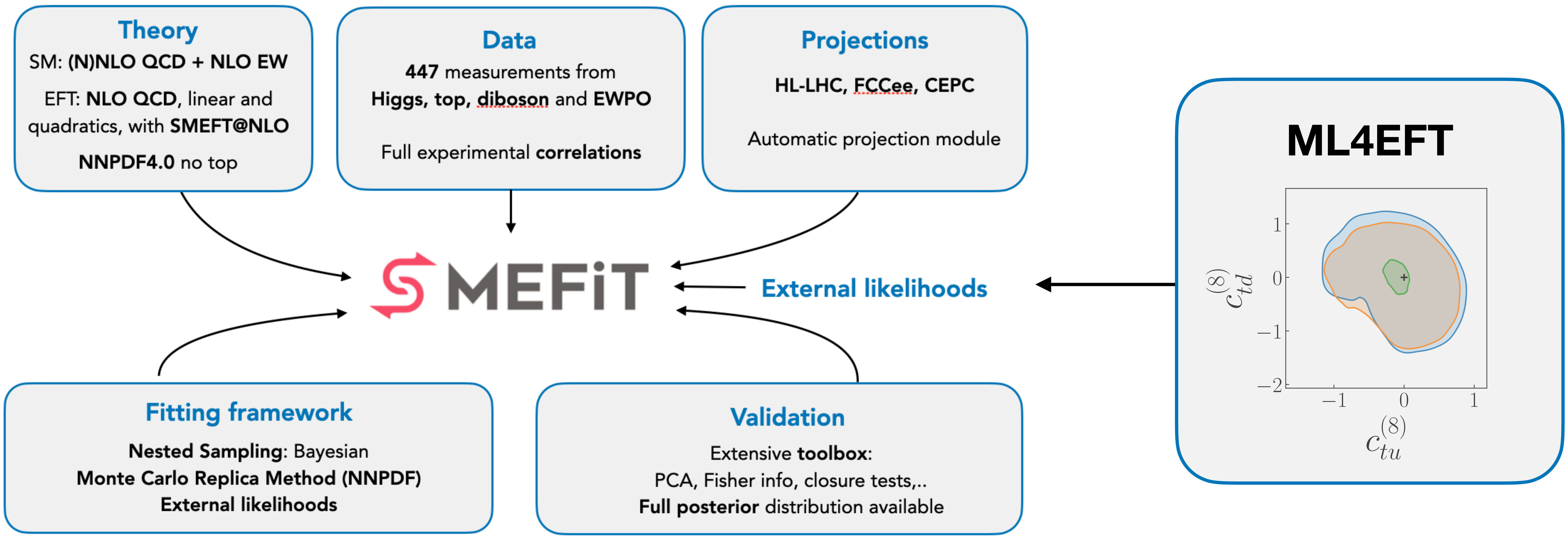
$$p_T^{\ell\bar{\ell}} \in [0, 10, 20, 40, 60, 100, 150, 400, \infty) \text{ GeV},$$
$$\eta^\ell \in [0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.5] .$$

- **Binned vs unbinned** in $(p_T^{\ell\bar{\ell}}, \eta_\ell)$ small improvement relative to binned setup
- **2 features vs 18 features:** big increase in sensitivity

ML4EFT + SMEFiT

A combined framework

Best of two worlds?



- The ultimate global EFT fit combines **binned** and **multivariate unbinned** ML observables
- We need a framework that connects them

A combined framework

The SMEFiT (binned) likelihood

$$-2\log \mathscr{L}(c) = \frac{1}{n_{\mathrm{dat}}} \sum_{i,j=1}^{n_{\mathrm{dat}}} \left(\sigma_{i,\mathrm{SMEFT}}(c) - \sigma_{i,\mathrm{exp}}\right) \left(\mathrm{cov}^{-1}\right)_{ij} \left(\sigma_{j,\mathrm{SMEFT}}(c) - \sigma_{j,\mathrm{exp}}\right)$$

And the multivariate unbinned ML likelihood (ratio)

$$\mathcal{L}(\boldsymbol{c}) = \frac{\nu_{\mathrm{tot}}(\boldsymbol{c})^{N_{\mathrm{ev}}}}{N_{\mathrm{ev}}!} e^{-\nu_{\mathrm{tot}}(\boldsymbol{c})} \prod_{i=1}^{N_{\mathrm{ev}}} f_\sigma\left(\boldsymbol{x}_i, \boldsymbol{c}\right)$$

Together form a powerful overall likelihood

$$\log\mathcal{L}(c) = \sum_{k=1}^{N_D^{(\mathrm{unbinned})}} \log\mathcal{L}_k^{\mathrm{unbinned}}(c) + \sum_{k=1}^{N_D^{(\mathrm{binned})}} \log\mathcal{L}_k^{\mathrm{binned}}(c)$$

A toy setup

SMEFiT

- ATLAS_CMS_SSinc_RunI
- ATLAS_SSinc_RunII
- CMS_SSinc_RunII

- ATLAS_WW_13TeV_2016_memu
- ATLAS_WZ_13TeV_2016_mTWZ
- CMS_WZ_13TeV_2016_pTZ
- CMS_WZ_13TeV_2022_pTZ

- ATLAS_STXS_runII_13TeV
- ATLAS_WH_Hbb_13TeV
- ATLAS_ZH_Hbb_13TeV
- ATLAS_ggF_13TeV_2015
- CMS_ggF_aa_13TeV
- ATLAS_ggF_ZZ_13TeV
- CMS_H_13TeV_2015_pTH

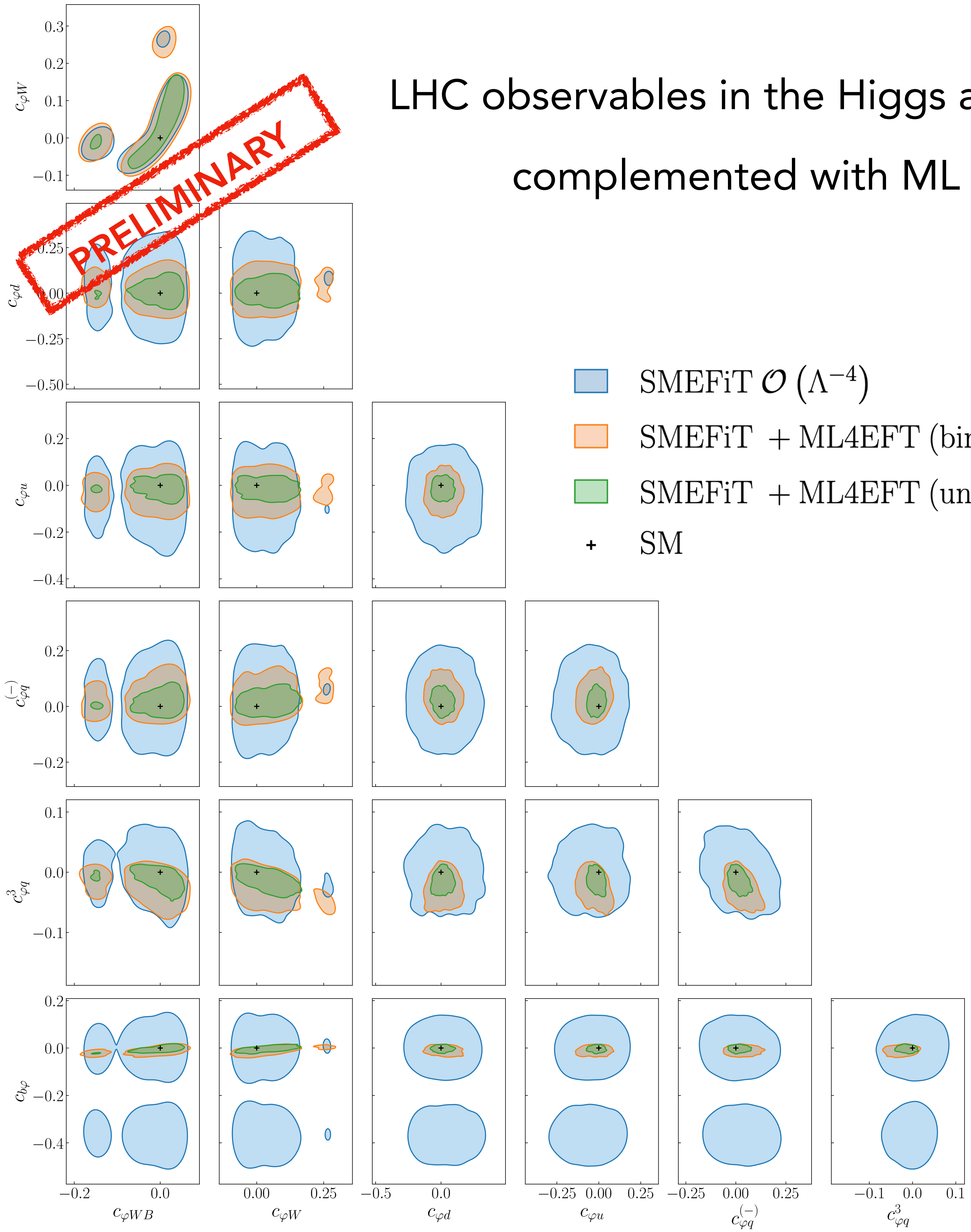
+

ML4EFT Run III projections

ATLAS_ZH_Hbb_13TeV

LHC observables in the Higgs and diboson sector....

complemented with ML observables



SMEFiT $\mathcal{O}\left(\Lambda^{-4}\right)$

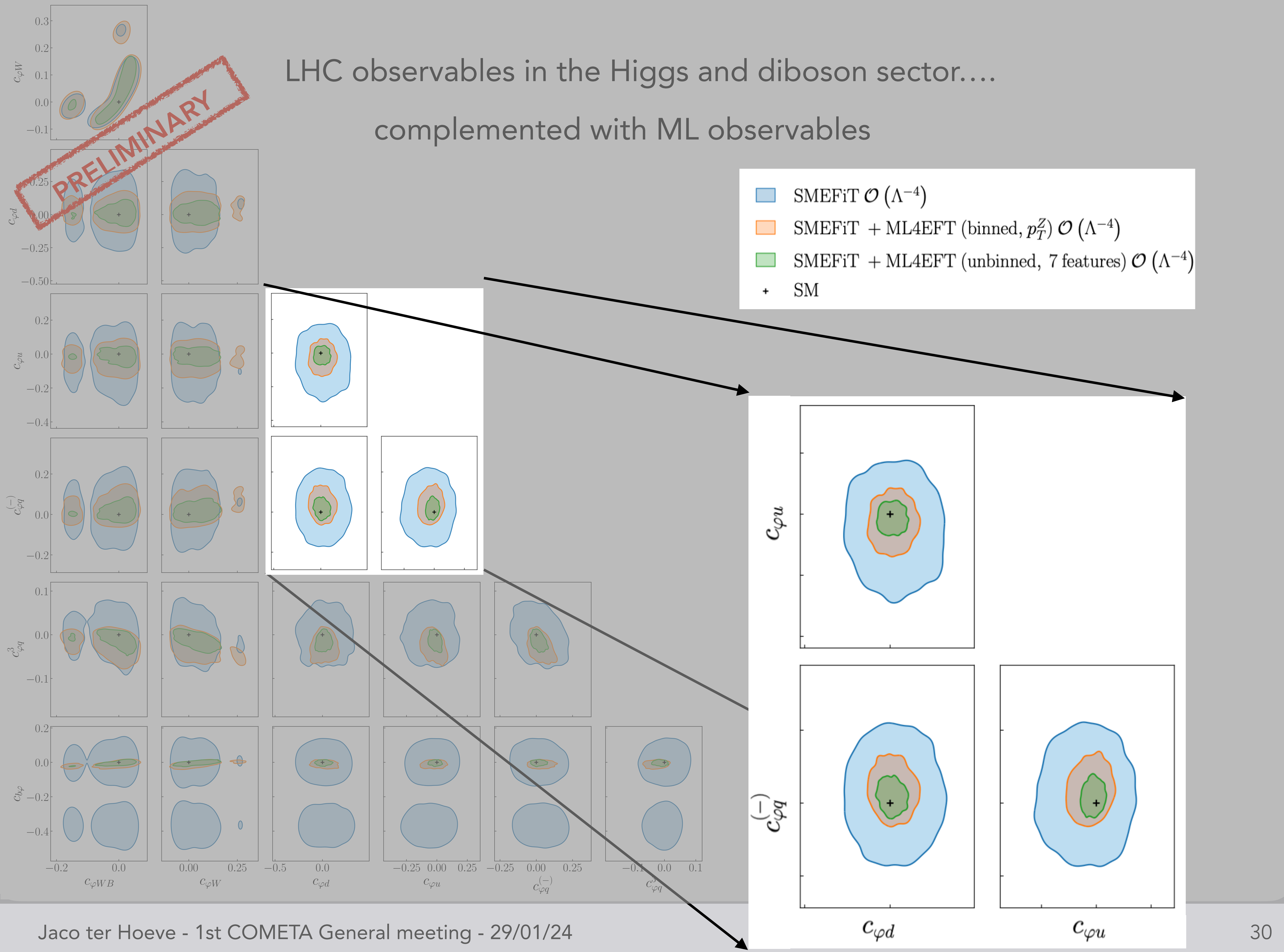
SMEFiT + ML4EFT (binned, p_T^Z) $\mathcal{O}\left(\Lambda^{-4}\right)$

SMEFiT + ML4EFT (unbinned, 7 features) $\mathcal{O}\left(\Lambda^{-4}\right)$

+ SM

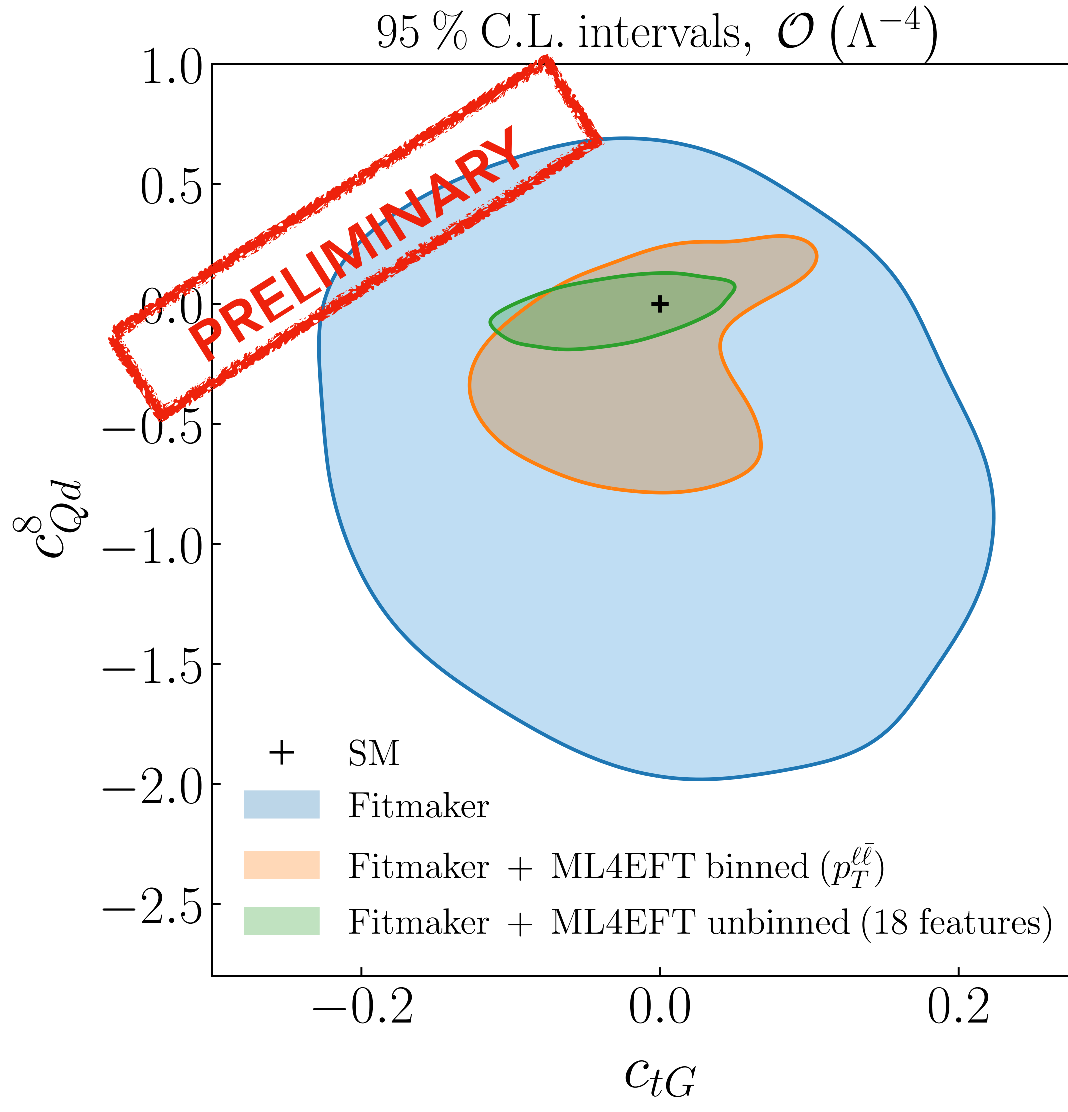
LHC observables in the Higgs and diboson sector….

complemented with ML observables



A combined framework

- Similar picture in the top sector



Summary

- State of the art global SMEFT fit with the latest run II results
- ML4EFT integrates unbinned multivariate observables into global SMEFT fits with a faithful uncertainty estimate through the replica method
- Case study in the Higgs and top sector
- Global EFT fits show increased sensitivity to unbinned observables
- Please visit **`ML4EFT`** on GitHub (documentation + **tutorial**)

lhcfitnikhef.github.io/ML4EFT

Thank you!